

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 06:47:02 ; Search time 2892.64 Seconds  
(without alignments)  
2958.221 Million cell updates/sec

Title: US-09-503-596-2  
Perfect score: 634  
Sequence: 1 ggaattccaggagggtgcag.....ataacttttttagatttag 634

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	607.6	95.8	628	10	BG286487
2	601.4	94.9	635	9	AA843663
3	598.6	94.4	636	10	BG570228
4	598.4	94.4	632	10	BG288820
5	597.4	94.2	620	9	AT635512
6	597.2	94.2	903	10	BG434107
7	596.4	94.1	630	9	AT652163
8	596	94.0	635	9	AT651387
9	590.4	93.1	637	9	AL047557
10	588.8	92.9	678	10	BG287253
11	586	92.4	616	9	AW237388
12	585	92.3	900	10	BG261014
13	584.4	92.2	632	10	BG289969
14	584.2	92.1	608	10	BE379934
15	582.8	91.9	654	10	BG287532
16	581.6	91.7	886	10	BG622132
17	577	91.0	905	10	BG572115

18	573.2	90.4	633	10	BG290820
19	561.8	88.6	682	10	BG261077
20	558	88.0	612	10	BG288203
21	556.6	87.8	597	9	AW236921
22	553	87.2	589	10	BE856868
23	550.4	86.8	591	9	AA732844
24	544.8	85.9	776	10	BI821655
25	542.8	85.6	576	9	AA777211
26	537.6	84.8	554	10	BE549708
27	528.4	83.3	570	9	AL047558
28	524	82.6	571	9	AI192168
29	521.8	82.3	559	9	AW051840
30	517	81.5	545	10	BF434751
31	515.8	81.4	767	10	BG289936
32	514.6	81.2	777	10	BF027862
33	513.6	81.0	621	10	BG285906
34	512.6	80.9	539	10	BG655515
35	511	80.6	560	9	AA861653
36	509	80.3	527	10	BG542023
37	508	80.1	542	10	BF590882
38	507.8	80.1	548	9	AI194006
39	500.4	78.9	619	10	BG285965
40	493	77.8	559	10	BE566213
41	492.8	77.7	505	9	AW631118
42	491.8	77.6	505	9	AT1767352
43	491.6	77.5	524	9	AW241575
44	491.4	77.5	855	10	BE568599
45	488.8	77.1	606	9	AA126737

## ALIGNMENTS

RESULT 1

LOCUS BG286487 628 bp mRNA linear EST 21-FEB-2001  
DEFINITION 602382975F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4500475 5', mRNA sequence.

ACCESSION BG286487

VERSION BG286487.1 GI:13039399

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 628)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Cloned by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10366 row: e column: 20

High quality sequence stop: 610.

FEATURES

source

1..628  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4500475"  
/clone.lib="NIH\_MGC\_93"  
/tissue.type="transitional cell papilloma, cell line"  
/lab.host="DH10B (phage-resistant)"  
/note="Organ: bladder; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dt primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

```
BASE COUNT      208 a   107 c   141 g   172 t
ORIGIN

Query Match      95.8%; Score 607.6; DB 10; Length 628;
Best Local Similarity 99.3%; Pred. No. 3e-153;
Matches 610; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 20 gcttctctcaccctgaagaataatcctcagaaactcacaaaactgctgctcttctgt 79
Db 1 GCTTCCTCTCACCTTGGAAGAAATATCTAGAAACTCAGAAATGCTGTGATGCTTTTGT 60

Qy 80 aggtacctggaacttctcctcagtgaaactttgatgattatgaagaagttaggagt 139
Db 61 AGGTACCTGGAAACTTGTCTCCAGTGAAGAACTTTGATGATTATGAAAGAGTAGGAGT 120

Qy 140 gggcttgcaccaggaagtggtgcgatggccaaactaacatgatcatcagtgtaa 199
Db 121 GGGCTTTGCCACCAGGAAGTGGCTGGCATGGCCAAACTTAACTGATCATCAGTGTGAA 180

Qy 200 tgggagtgatcacaccattaaatctgaaagtcaccttttaaaatactgagattcctcat 259
Db 181 TGGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAATACTAGAGATTTCCTTCAT 240

Qy 260 actgggccaggaatttgacgaagtcactgcagatgacaggaagtcaagagcacataac 319
Db 241 ACTGGCCAGGAATTTGACGAAGTCACTGTCAGATGACAGGAAAGTCAAGAGCAGCATAAC 300

Qy 320 cttagatgggggtgctcctgtacatgacagaatggatgaaatacaacacacataaa 379
Db 301 CTTAGATGGGGGTGCTCCGTGACATGTCAGAAATGGATGGAATAACCAACCATATAA 360

Qy 380 gagaaacagagagatgataaactggtggtgaatgctcatgaaagcgctcactccac 439
Db 361 GAGAAACAGAGGATGATAAATGTTGGTGAATGCGTCATGAAAGCGTCACTTCAC 420

Qy 440 gagagtttatgagagagacataagccaagggagcgttgacctggaactggaagtcgcatgaa 499
Db 421 GAGAGTTTATGAGAGAGCATAGCAAGGAGCGTTGACCTGGACTGAAGTTGCGATTGAA 480

Qy 500 ctctacaacatctgtgggatatattgtcctcagaaagatatgttttccctcatttag 559
Db 481 CTCTACACATCTGTGGGATATATGTTCAAAAGATATTTGTTTTCATGATTTAG 540

Qy 560 caagcaagtaatttctcccaagctgattttattcaatatgtttacgttggtggttaataac 619
Db 541 CAAGCAACTAATTTCTCCCAAGCTGATTTTATTCAATATGTTAGCTGTGTTAATAAAC 600

Qy 620 ttttttagattta 633
Db 601 TTTTATGATTTA 614

RESULT 2
AA843663/c
LOCUS
DEFINITION
ak08g10.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
IMAGE:1405410.3' similar to gb:J02874 FATTY ACID-BINDING PROTEIN,
ADIPOCYTE (HUMAN);, mRNA sequence.
ACCESSION
AA843663
VERSION
AA843663.1 GI:2930181
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 635)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-1@mail.nih.gov
```

```
cdna Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cdna Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 681 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 472.
Location/Qualifiers
1. 635
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1405410"
/clone_lib="Soares_parathyroid_tumor_NbHPA"
/tissue_type="parathyroid tumor"
/dev_stage="adult"
/note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
) with a modified polylinker; Site.1: Not I; Site.2: Eco
RI; 1st strand cdna was primed with a Not I - oligo(dT)
primer
[5'-TGTTACCAATCTCAAGTGGAGCGCGGCACCAATTTTCTTTTCTTTTCTTTT
TTTTT-3'], double-stranded cdna was size selected, ligated
to Eco RI adapters (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of a modified pT7T3
vector (Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
adenomas was kindly provided by Dr. Stephen Marx, National
Institute of Diabetes and Digestive and Kidney Diseases,
NIH."
```

BASE COUNT 174 a 142 c 106 g 213 t  
ORIGIN

Query Match 94.9%; Score 601.4; DB 9; Length 635;  
Best Local Similarity 99.0%; Pred. No. 1.4e-151;  
Matches 605; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 16 tgcagcttccttcacacctgaagaataatcctcagaaactcaacaaatgtgtgattgt 75  
Db 635 TGCAGCTTCCTTCACCTTGAAGAATAATCTTGAAACTCACAAGATGTGTGATGCTT 576

Qy 76 tttagtgacctggaacctgtctccagtgaaacctttgatgattatgaaagaagtag 135  
Db 575 TTGTAGGTACCTGGAAACTTGTCTCCAGTGAAACTTTGTATGATGATTTATGAAGAAGTAG 516

Qy 136 gagtgggcttggccaccaggaaagtggctggcatggcgccatggccacaaaccttaacatgatcatcagtg 195  
Db 515 GAGTTGGCTTTGCCACCAGGAAGTGGCTGGCATGGCCAAACCTTAACATGATCATCAGTG 456

Qy 196 tgaatgggagtgatgataccattaaatctgaaagtacaccttttaaaatactagatttct 255  
Db 455 TGAATGGGAGTGTGATCACCATTAAATCTGAAAGTACCTTTAAATAATACTGAGATTTCT 396

Qy 256 tcatactggccagaaatttgcacaaagtcactgcagatgcagaaagtcagaaagtcagagcaca 315  
Db 395 TCATACTGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAGTCAAGAGCAGCA 336

Qy 316 taaccttagatgggggtgctcctggtacatgtgcagaaatggatggaaatacaaccacca 375  
Db 335 TAACCTTAGATGGGGTCTCCTGGTACATGTGCAGAAATGGATGGAAATAACACCACCA 276

Qy 376 taaagaaaaacgagagatgataaactggtggtggaatgcgtcatgaaagcgctcatt 435  
Db 275 TAAAGAAAAACGAGAGATGATAAATCTGTTGGTGAATGCTCATGAAGCGCTCACTT 216

Qy 436 ccacgagatttatgagagacataaaccaggaagcgttgacctggactgaagttcgat 495  
Db 215 CCACGAGATTTATGAGAGACATAAGCAAGGACGTTTGACCTGGACTGAAGTTGGAT 156

QY 496 tgaactacaacattctgtgggatatattgttccaaaagatatgttgttttccctgat 555  
 Db 155 TGAACCTACAACATCTGTGGGATATTGTTCAAAAAGATATTGTTGTTTCCATGAT 96  
 QY 556 ttagcaagcaagtaattttcccaagctgatatttattcaatatggttacgtggttaaa 615  
 Db 95 TTAGCAAGCAACTAATTTTCTCCCAAGCTGATTTTATTCATATGTTACGTTGGTTAAA 36  
 QY 616 taactttttt 626  
 Db 35 TAAACTTTTT 25

RESULT 3  
 BG570228 636 bp mRNA linear EST 10-APR-2001  
 LOCUS 602590629F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4717518 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG570228  
 VERSION BG570228.1 GI:13577881  
 KEYWORDS EST.  
 RCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 636)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1567 row: i column: 07  
 High quality sequence stop: 632.  
 Location/Qualifiers  
 1. .636  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4717518"  
 /clone\_lib="NIH\_MGC\_77"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatgcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGAGCGCGCATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

BASE COUNT 206 a 114 c 144 g 172 t  
 ORIGIN

Query Match 94.4%; Score 598.6; DB 10; Length 636;  
 Best Local Similarity 98.4%; Pred. No. 8.2e-151;  
 Matches 615; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 6 tccaggagggtgcagcttccttcaccttgaagaataatcctagaaaactcacaaaatg 65  
 Db 10 TCCTGAACATGACGCTTCCTCTCACCTTGAAGAATAATCTGAAACATCAACAAATG 69  
 QY 66 tgtgatctttgttagtgacctggaactgtctccagtgaaaacttgatgatattatg 125  
 Db 70 TGTGATGCTTTGTAGTACCTGGAACACTTGTCTCCAGTGAAACTTTGATGATTATG 129

QY 126 aaagaaagtagagtgaggcttttccaccagagaaagtgcctggcattgccaacacctaactg 185  
 Db 130 AAAGAAGTAGAGTGGGCTTTTCCACCAGGAAAGTGGCTGGCATGGCCAAACCTAACATG 189  
 QY 186 atcatcagtggaatgggagtgatcaccattaaactgaaagtacacctttaaataact 245  
 Db 190 ATCATCAGTGTGAATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTTAAANACT 249  
 QY 246 gagatttccttcatactgggcccaggaatttgacgaagtacactgcagatgcagagaaagtc 305  
 Db 250 GAGATTTCTTCACTTGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAAGTC 309  
 QY 306 aagagcaccataaccttagatgggggtgctcctgtacatgtgcagaaatgggagtggaata 365  
 Db 310 AAGAGCACCATAACCTTAGATGGGGGTCTCCTGTACATGTGCAGAAATGGGATGGAAAA 369  
 QY 366 tcaaccaccataaagagaaacagagagatataaactggtggtggaatgcgtcatgaaa 425  
 Db 370 TCAACCACTAAAGAGAAACAGAGAGATGATAAACTGGTGGTGAATGGCTCATGAAA 429  
 QY 426 gggtcacttcacgagagtgattatgagagagcataagccaaaggagcgttgacctggactg 485  
 Db 430 GGCCTCACTTCCACGAGAGTTTATGAGAGAGCATAAGCCAAAGGACGTTGACCTGGACTG 489  
 QY 486 aagttccattgaactcaacattctgtggatatattgttcaaaaagattgttgt 545  
 Db 490 AAGTTCGATTGAACTCTACAACTTCTGTGGATATATTGTTCAAAAAGATATGTTGT 549  
 QY 546 ttccctgattagcaagcaagtaatttctcccaagctgattttattcaatatggttac 605  
 Db 550 TTTCATGATTTAGCAAGCAACTAATTTCTCCCAAGCTGATTTTATTC-ATATGGTTAC 608  
 QY 606 gttggttaataacttttttagat 630  
 Db 609 GTTGGTTAAATAAACTTTTGTAGAT 633

RESULT 4  
 BG288820 632 bp mRNA linear EST 21-FEB-2001  
 LOCUS 602388151F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4516981 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG288820  
 VERSION BG288820.1 GI:13044042  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 632)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHAM10409 row: e column: 14  
 High quality sequence stop: 604.  
 Location/Qualifiers  
 1. .632  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4516981"  
 /clone\_lib="NIH\_MGC\_93"  
 /tissue\_type="transitional cell papilloma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: bladder; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: Sali; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 208 a 112 c 142 g 170 t  
ORIGIN  
Query Match 94.4%; Score 598.4; DB 10; Length 632;  
Best Local Similarity 99.0%; Pred. No. 9.2e-151;  
Matches 602; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 19 agctctcttcacactgaagaatactcctagaagaactcacaataatgtgtgatctttg 78  
Db 1 AGCTTCCTCTCCTCCTGGAAGTAATCCTAGAAAACCTCACAATAATGTGTGATCTTTG 60  
Qy 79 tagctactggaactgtctccagtgaaacttttgatgattatgaagaagtagag 138  
Db 61 TAGGTACTCTGGAACCTTGTCTCCAGTGAAACCTTTGTGATGATTATGAAAGAGTAGGAG 120  
Qy 139 tgggtcttgcaccaggaagtgtgctggcagtgccaaacccaacatacatgatcatcagtgta 198  
Db 121 TGGGCTTTGCCACCAGGAAGTGTCTGGCATGGCCAAACCTAACATGATCATCATGTGA 180  
Qy 199 atggggatgtgatacaccattaaatctgaaagtacactttaaaatactgagatttccttca 258  
Db 181 ATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAAAATACCTGAGATTTCTCTCA 240  
Qy 259 tactggccaggaatttgacgaagtcactgcagatgacagagaaagtcaagagaccataa 318  
Db 241 TACTGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAAGTCAAGAGCACCATAA 300  
Qy 319 ccttagatgggggtgtcctgtgacatgtgcagaaatgggtggaataatcaaccaccataa 378  
Db 301 CCTTAGATGGGGTGCTCTGGTACATGTGCAGAAATGGGATGGGAATCAACACCATAA 360  
Qy 379 agagaacagagagatgataaaactggtgtggaatgcgtatgaaaggcgtcactcca 438  
Db 361 AGAAAAACAGAGAGATGATAAATGCTGGTGGTGAATCGTCATGAAAGGCGTCACTTCCA 420  
Qy 439 cgaggtttatgagagacataagccaaagacgttgacactgactgaagttcgcattga 498  
Db 421 CGAGAGTTTATGAGAGAGATAAGCCAAAGGACGTTGACCTGGACTGAAGTTGCGATTGA 480  
Qy 499 actcaacaacttctgtggatattgttcaaaagatatgtgttttccctgattta 558  
Db 481 ACTCTACAACATCTGTGGGATATATGTTCAAAAGATATGTTGTTTCCATGATTTA 540  
Qy 559 gcaagaagtaattttctcccaagctgattttatccaatatggtacgttgggttaataa 618  
Db 541 GCAAGCAACTAATTTTCTCCCAAGCTGATTTTATTCATATGTTTACGTTGGTTAAATAA 600  
Qy 619 cttttttt 626  
Db 601 ACTTTCT 608

RESULT 5  
LOCUS AI635512/c 620 bp mRNA linear EST 14-DEC-1999  
DEFINITION t895a09.x1 NCI\_CGAP\_G6 Homo sapiens cDNA clone IMAGE:2239000 3'  
similar to gb:J02874 FATTY ACID-BINDING PROTEIN, ADIPOCYTE (HUMAN  
); mRNA sequence.  
ACCESSION AI635512  
VERSION AI635512.1 GI:4686842  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 620)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

# Tumor Gene Index

## JOURNAL COMMENT

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 726 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 445.

## FEATURES

### source

1. 620  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2239000"  
/clone\_lib="NCI\_CGAP\_G6"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH10B"

/notes="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA  
from the normalized library NCI\_CGAP\_G64 was prepared, and  
ss circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (cloneIDs  
1257096-1258631, 1469064-1470983, and 1475592-1476743).  
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 172 a 143 c 106 g 198 t 1 others

## ORIGIN

Query Match 94.2%; Score 597.4; DB 9; Length 620;  
Best Local Similarity 98.7%; Pred. No. 1.7e-150;  
Matches 612; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
Qy 16 tgcagctctcttcacactgaagaataactcctagaagaactcacaataatgtgtgactctt 75  
Db 620 TGCAGCTTCTCTTCACCTGAAAGAAATATCTCTGAAAGTCCAAAAATGTGTGATGCT 561  
Qy 76 ttgtagctactgagaactgtctccagtgaaactttgatgtatgatataaagaagtag 135  
Db 560 TTGTAGTACTGGAAGTGTCTCCAGTGAAACTTTGTATGATATATGAAAGAGTAG 501  
Qy 136 gagtgggctttgccaccaggaagtgctggcagtgcccaaaccttaacatgatcatcagtg 195  
Db 500 GAGTGGGCTTTGCCACCAGGAAGTGGCTGGCATGGCCAAACCTAACATGATCATCAGT 441  
Qy 196 tgaatggggatgtgatcaccattaaatcctgaaagtacatttaaaatactgagatttcc 255  
Db 440 TGAATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAAAATGAGATTTCT 381  
Qy 256 tcactactggccagggaatttgacgaagtcactgcagatgacagagaaagtcaagagaccca 315  
Db 380 TCATACTGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAAGTCAAGAGACCA 321  
Qy 316 taaccttagatgggggtgtcctgtacatgtgcagaatgggagtggaataatcaaccacca 375  
Db 320 TAACCTTAGATGGGGGTGTCTCTGTATACATGTGCAGAAATGGGATGGGAAATCAACCA 261  
Qy 376 taaagagaaacagagagatgataaaactggtgtggaatgcgtcatgaaaggcgtcactt 435  
Db 260 TAAAGAAAACAGAGAGATGATAAACTGTGTGGAAATCGTCATGAAAGGCGTCACTT 201  
Qy 436 ccacgaggtttatgagagagcataaagcgaagacgttgacctggactgaagtctgcac 495  
Db 200 CCACGAGAGTTTATGAGAGAGCATAAGCCAAAGGACGCTTGACCTGGACTGAAGTTCGCAT 141



```
Qy 496 tgaactacacattctgtgggatatttccaaaaagattgtgttttccctgat 555
Db 140 TCAACTCTACAACTCTGTGGGATATATTGTTCAAAAAGATATTGTTGTTTCCATGAT 81
Qy 556 ttagcaagcaagtaattctcccaagctgatttattcaatatggttacgtgttaaa 615
Db 80 TTAGCAAGCACTAAATTTCTCCCAAGCTGATTTATTTCAATATGTTACGTTGGTTAAA 21
Qy 616 t-aacttttttagatttag 634
Db 20 TAAACTTTTTTAGATTAG 1

RESULT 6
BG434107 903 bp mRNA linear EST 14-MAR-2001
LOCUS 602506025F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4603278 5',
DEFINITION mRNA sequence.
ACCESSION BG434107
VERSION BG434107.1 GI:13340613
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgaps-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M337 row: a column: 07
High quality sequence stop: 624.
FEATURES
source 1. 903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4603278"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggcgctcgcc); Site 2: SfII (ggcattatggc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dT(30)BN-3'
(sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dT(30)BN-3'
(where B = A, C, G, or N = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 276 a 204 c 217 g 206 t
ORIGIN

Query Match 94.2%; Score 597.2; DB 10; Length 903;
Best Local Similarity 97.1%; Pred. No. 2.le-150;
Matches 608; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 6 tccagagggtgcagcttccttcacattgaagaataatcctagaaactcacaaaatg 65
Db 11 TCCGTAAACATGCGAGCTTCCTTCACCTTGAAGAATAATCTCTGAAACATCACAAAATG 70
Qy 66 tgtgatgctttgttaggtaccttggaacttgtctccagtgaaactttgtgatattatg 125
Db 71 TGTGATGCTTTTGTAGGTACTGGAACACTTCTCTCCAGTGAACACTTTGATGATATATG 130
```

```
Qy 126 aagaagtaggagtggtgttggccaccaggaaagtggctgcatggccaaacctaacatg 185
Db 131 AAAGAAGTAGGAGTGGGCTTTGGCCACCAGGAAAGTGGCTGGCATGCCCAACCTAACATG 190
Qy 186 atcatcagtgaatggggatgtgataccattaaatgaagacagacattttaaataact 245
Db 191 ATCATCATGTGAATGGGGATGTGATCACCATTAAATCTGAAGTACCTTTTAAAAATACT 250
Qy 246 gagatttcttcatactggccaggaaatttgacgaagtcactgcagatgacaggaagtc 305
Db 251 CAGATTTCTTCACTGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAGTC 310
Qy 306 aagagcaccataaccttagatgggggtgtcctgtgtacatgtgcagaatgggatgaaaa 365
Db 311 AAGAGCACCATAAACCTTAGATGGGGGTGCTCCTGTACATGTGCAGAAATGGGATGAAAA 370
Qy 366 tcaaccaccataaagagaaaaacgagagatgataaactgggtggaatgcgtcatgaaa 425
Db 371 TCAACCACCATAAAGAGAAAACGAGAGATGATAAATCTGGGTGGATGCGTCATGAAA 430
Qy 426 ggcgtcactccacgagagatttatgagagacataagccaaagggagcttgacctggactg 485
Db 431 GCGGTCACTTCCACGAGAGTATTATGAGAGAGCATAAAGCAAGGGACGTTGACCTGGACTG 490
Qy 486 aagttcgattgaactctacaacattctgtgggatataattgtcacaagaatattgttgt 545
Db 491 AAGTTCGCATTGAACCTTACAAACATTTCTGTGGGATATATGTTTCACACAGATATTGTCT 550
Qy 546 ttccctgatttagcaagcaagtaatttctcccaagctgatttattcaatatggttac 605
Db 551 TTTCCATGATTAGACAGCAACTAATTTCTCCCAAGCTGATTTTATTCATATGTTTAC 610
Qy 606 gttggttaataaacttttttagatt 631
Db 611 GTTGGTTAATAAACTTTCTCTAGATT 636

RESULT 7
AI652163 630 bp mRNA linear EST 17-DEC-1999
LOCUS wb47g01.x1 NCI-CGAP GC6 Homo sapiens cDNA clone IMAGE:2308848 3',
DEFINITION similar to gb:J02874 FATTY ACID-BINDING PROTEIN, ADIPOCYTE (HUMAN
); mRNA sequence.
ACCESSION AI652163
VERSION AI652163.1 GI:4736142
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 630)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 717 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 404.
FEATURES
source 1. 630
/organism="Homo sapiens"
```



```

Db 331 CCTTAGATGGGGTGCTCTGTACATGTGCAGAAATGGGATGGAAATCAACCCCATAA 272
Qy 379 agagaaacagagagatgataaactggtggtggaatcgctcatgaaaggcgcactcca 438
Db 271 AGAGAAAACGAGAGATGATAAATGGTGTGGTGAATGGTCAATGAAGGGCGTCACTTCCA 212
Qy 439 cgagagttatgagagacataaagcgaagagacgttgacccggaactgaagtcgacattga 498
Db 211 CGAGAGTTTATGAGAGAGCATAAACCAAGGAGACCTTGACCTGGGACTGAAGTTCGCATTGA 152
Qy 499 actcacaaactctgtgggatatattgttcaaaaagatatgttgtttccctcgattta 558
Db 151 ACTCTAACAACTCTGTGGGATATATTGTTCAAAAAGATATTGTTGTTTCCATGATTTA 92
Qy 559 gcaagcgaagtaattttctccaaagctgattttatccaatggttacgttggttaaat-a 617
Db 91 GCAAGCAACTAAATTTCTCCCAAGCTGATTTTATTAATATGTTACGTTGGTTAAATAA 32
Qy 618 acttttttagatttag 634
Db 31 ACTTTTTTAGATTAG 15

RESULT 9
AL047557 637 bp mRNA linear EST 01-MAR-2000
LOCUS DKFZ586E0221_r1 586 (synonym: hutel) Homo sapiens cDNA clone
DEFINITION DKFZ586E0221, mRNA sequence.
ACCESSION AL047557
VERSION AL047557.3 GI:5936382
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ottewaeldeider, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Ottewaeldeider, et al.)
Unpublished (1999)
On Apr 30, 1999 this sequence version replaced gi:5866815.
Contact: Ottewaeldeider B
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. s1 sequence
also available
This clone (DKFZ586E0221) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..637
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZ586E0221"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
214 a 106 c 146 g 171 t
BASE COUNT 93.18; Score 590.4; DB 9; Length 637;
ORIGIN Best Local Similarity 98.9%; Pred. No. 1.3e-148;
Matches 605; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 16 tgcagcttccttcaccttgaagaataatcctagaaaactcacaaaatgtgtgatgctt 75
|||||

```

```

Db 1 TGCAGCTTCCTTCTCACCTTGAAGAAATAACTCTAGAAAACTCACAAAATGTGTGATGCTT 60
Qy 76 ttgtaggtacgtgaaactgtctccagtgaacttttgatgatttatatgaagaagttag 135
Db 61 TTGTAGGTACCTGGAAACTTGTCTCCAGTAAAACTTTTGTATGATTATATGAAGAAGTAG 120
Qy 136 gagtgggctttgcacacaggaagtggctggcatggcacaacctaaacatgatcatcagt 195
Db 121 GAGTGGGCTTTGCCACCAGGAAGTGGCTGGCATGGCCAAAACCTTAACATGATCATCAGTG 180
Qy 196 tgaatgggagtgatcacaccattaaactgaaagtacacctttaaataactgagatttct 255
Db 181 TGAATGGGGATGTGATCACCATTAACTCTGAAGTACCTTTAAAAAATACTGAGATTTCCT 240
Qy 256 tcatactgggcccaggaatttgacgaagtcaactgcagatgcacaggaaaagtcaagagacaca 315
Db 241 TCATACTGGGCCAGGAAATTTGACGAAGTCACTGCGAGATGACAGGAAAGTCAAGAGACCCA 300
Qy 316 taacctagatggggtgctcctggtacatgtgcagaatgggatggaatacaacacaca 375
Db 301 TAACCTTAGATGGGGGTGCTCTGTGATCATGTGCAGAAATGGGATGGAAATCAACACCA 360
Qy 376 taaagagaaaaacagagagatgataaaactgggtgggaatgcgtcatgaaagcgctcact 435
Db 361 TAAAGAGAAAAACGAGAGGATGATAAAGTGGTGGTGAATGCGTCATGAAAGCGTCACCT 420
Qy 436 ccagagaggtttatgagagacataaagccaaaggagcgttgacctggaactgaagtcgat 495
Db 421 CCACGAGAGTTTATGAGAGAGCATAAAGCAAGGAGCTTGACCTGGACTGAAGTTCCGAT 480
Qy 496 tgaactctacaacattctgtggatatattgttcaaaaagatatgtgttttccctgat 555
Db 481 TGAACCTACAACTATCTGTGGGATATATTGTCAAAAGATATTGGTGTGTTTCCATGAT 540
Qy 556 ttgacaaagcaagtaattttctccaaagctgattttattcaaat-ggttacgttggttaa 614
Db 541 TTAGCAACCACTAATTTCTCCCAAGCTGATTTATTCAATATGGTTACGTTGGTTAA 600
Qy 615 ataactttttt 626
Db 601 ATAACTTTTTT 612

RESULT 10
BG287253 678 bp mRNA linear EST 21-FEB-2001
LOCUS 602381981P1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499252 5',
DEFINITION mRNA sequence.
ACCESSION BG287253
VERSION BG287253.1 GI:13040909
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L1AM10363 row: b column: 21
High quality sequence stop: 625.
FEATURES
source
1..678
/organism="Homo sapiens"

```

```
/db_xref="taxon:9606"  
/clone="IMAGE:4499252"  
/clone_lib="NIH_MGC_93"  
/tissue_type="transitional cell papilloma, cell line"  
/lab_host="DH10B (phage-resistant)"  
/notes="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;  
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH_MGC Library."
```

BASE COUNT 225 a 130 c 152 g 171 t

Query Match 92.9%; Score 588.8; DB 10; Length 678;  
Best Local Similarity 98.7%; Pred. No. 3.6e-148;  
Matches 604; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```
16 tgcagcttccttcacattgaaataataactcctagaaactcacaataatgtgtatgctt 75  
|||||  
19 TGCAGCTTCCTTCACCTTGAAGATAATCTCTAGAAACTCACAATAATGTGTATGCTT 78  
|||||  
76 ttgtagtagtgcgaactgtctccagtgaactcttgatgattatataaagaagtag 135  
|||||  
79 TTGTAGGTACCTGGAACTTGTCTCCAGTGAATACTTTGATGATTATATGAAGAAGTAG 138  
|||||  
136 gagtggccttggcaccaggaagtggctggcagtcgcaaacctcaacatgatcatcagt 195  
|||||  
139 GAGTGGGCTTTGCCACGAGAAAGTGGCTGGCATGGCCAAACCTTACATGATCATCAGTG 198  
|||||  
196 tgaatggggatgtgatcaccattaaatctgaagtacattctaaataactgagattcct 255  
|||||  
199 TGAATGGGATGTGATCACCATTAAATCTGAAGTACCTTTAAAAATACTGAGATTTCCT 258  
|||||  
256 tcatactggccaggaatttgacaaagtctcactgcagatgacagaaagtcaagaccaca 315  
|||||  
259 TCATACTGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGAAAGTCAAGACACCA 318  
|||||  
316 taaccttagatgggggtgtcctggtacatgtgcagaatggatggaaatcaaccacca 375  
|||||  
319 TAACCTTAGATGGGGTGTCTGTACATGTGCAGAAATGGATGGAAATCAACCACCA 378  
|||||  
376 taaagaaacagagagatgataaactgtgtggaatgcgtcatgaaaggcgtcaact 435  
|||||  
379 TAAAGAAACAGAGAGATGATAAATCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 438  
|||||  
436 ccagagagttatgagagacataagccaaggaacgttgacctggactgaagtgcgat 495  
|||||  
439 CCAGAGAGTTTATGAGAGAGCATAAAGCCAGGACGTTGACCTGGACTGAAGTTCGCAT 498  
|||||  
496 tgaactctacaactctgtgggatatattgttcaaaaagatatgtgttttccctgat 555  
|||||  
499 TGAACCTACAACTTGTGGGATATATGTTCAAAAGATATCGTCGTTTCCATGAT 558  
|||||  
556 ttagaagcaagttaatttctccc-aaagctgattttatcaatatggttacgttggttaa 614  
|||||  
559 TTAGCAAGCAACTAATTTCTCCCAAGAGCTGATTTTATCAATATGTTAGCGTGGTTAA 618  
|||||  
615 ataactttttt 626  
|||||  
619 ATAACTTTTTT 630  
|||||
```

RESULT 11  
AW237388/c  
LOCUS  
DEFINITION  
xnm71507.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2689621 3'-  
similar to gb:J02874 FATTY ACID-BINDING PROTEIN, ADIPOCYTE (HUMAN  
); mRNA sequence.  
ACCESSION  
VERSION  
AW237388  
AW237388.1 GI:6569777  
KEYWORDS  
EST.  
human.  
SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 616)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40UP from Gibco  
High quality sequence stop: 450.  
FEATURES  
Location/Qualifiers  
source 1..616  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2689621"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI\_CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 132376-132391, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 172 a 141 c 106 g 197 t

Query Match 92.4%; Score 586; DB 9; Length 616;  
Best Local Similarity 99.0%; Pred. No. 2e-147;  
Matches 600; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
Qy 22 ttccttccacttgaag-aataatcctagaaaactcacaataatgtgtgcttttga 80  
|||||  
Db 613 TCCCTTTCACCTTGAGAAATAATCTCTAGAAACTCACAATAATGTGTGATGCTTTTCTGA 554  
|||||  
Qy 81 ggtacctggaaactgtctccagtgaataacttgcattatgataatgaaagaagtaggagtg 140  
|||||  
Db 553 GGTACCTGGAAACTTGTCTCCAGTGAACACTTTTGATGATTATATCAAAAGAGTAGGAGTG 494  
|||||  
Qy 141 ggccttgcaccagaaagtgcctggcatggccaaacctcaatcatcatcagtgatgaat 200  
|||||  
Db 493 GCGTTTGCACACAGAAAGTGGCTGGCATGGCCAAACCTTAACATGATCATCAGTGTGAAT 434  
|||||  
Qy 201 ggggatgtgacacattcaatctgaaagtacacctttaaataactgagatttccctcata 260  
|||||  
Db 433 GGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAAAATACTGAGATTTCTTCATA 374  
|||||  
Qy 261 ctgggcccagggaatttgacgaagtctcagatgacaggaagtcagagagaccataacc 320  
|||||  
Db 373 CTGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAGTCAAGAGCACCATAACC 314  
|||||  
Qy 321 ttagatgggggtgctcctgggtacatgtgcagaaatgggagtgaaataaccaccataaag 380  
|||||  
Db 313 TTAGATGGGGTGTCTCGGTACATGTGCAGAAATGGGATGGGATGAAATCAACCACCAATAAG 254  
|||||  
Qy 381 agaaacagagagatgataaactggtggtggaatgcgcacatgaagcgctcactccacg 440  
|||||  
Db 253 AGAAACGAGAGGATGATAAATACTGCTGGTGGAAATGCGTCATGAAGGGCTCCTCCACG 194  
|||||
```

```

Qy 441 agagtttatgagagcataagcaagagcgttgacctggactgaagttcgcatggaac 500
Db 193 AGAGTTTATGAGAGCATAAGCCAAAGGACGTTGACCTGGACTGAAGTTCGCATTGAAC 134
Qy 501 tctcaacattctgtgggatatattgtccaaaagataattgttttccctgatttagc 560
Db 133 TCTCAACATCTGTGGGATATATGTTCAAAAGATATGTTGTTTCCATGATTAGC 74
Qy 561 aagcaagtaatttctcccaagctgattttatccaatatggttacgttggttaataact 620
Db 73 AAGCAACTAATTTCTCCCAAGCTGATTTTATCAATATGTTAGCTGGTTAAATAAAC 14
Qy 621 tttttt 626
Db 13 TTTTTT 8

RESULT 12
LOCUS BG261014 900 bp mRNA linear EST 13-FEB-2001
DEFINITION 602372632F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4480759 5',
mRNA sequence.
VERSION BG261014
KEYWORDS EST.
SOURCE BG261014.1 GI:12770830
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10314 row: p column: 08
High quality sequence start: 2
High quality sequence stop: 625.
Location/Qualifiers
1. 900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4480759"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/site="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 281 a 182 c 222 g 214 t 1 others
ORIGIN

Query Match 92.3%; Score 585; DB 10; Length 900;
Best Local Similarity 98.5%; Pred. No. 4.1e-147;
Matches 611; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

Qy 16 tgcagctctctctcacctgaagaataatcctagaataactcacaaaatgtgtgatgctt 75
Db 9 TGCAGCTTCTCTCTCACCTTGAGARATATCTCTAGAAACTCACAAAATGTGTGATGC-T 67
Qy 76 ttgtaggtacctggaacctgtctccagtgaaacctttgatgattatgatgaagaagtag 135
Db 68 TTGTAGGTACCTGGAACTTGCTCTCCAGTGAAACTTTGTGATGATTATATGAAAGAAGTAG 127

```

```

Qy 136 gagtgggttttgcaccaggaaagtgggtggcatggccaaacctaacatgatcatcagt 195
Db 128 GAGTGGGCTTTGCCACCAGGAAAGTGGTGGCATGCCAAACCTAACATGATCATCAGTG 187
Qy 196 tgaatgggagtgatcacaccattaaatctgaaagtcacctttaaaaatactagatttct 255
Db 188 TGAATGGGGATGTGATCACCAATTAATCTGAAAGTACCTTTAAAAATACTGAGATTCT 247
Qy 256 tcatactgggcccaggaatttgacgaagtcactgcagatgacagggaaagtcaagagacaca 315
Db 248 TCATACTGGGCCAGGAATTTGACGAAGTCACCTGCAGATGACAGGAAAGTCAAGAGACACCA 307
Qy 316 taaccttagatgggggtgtcctgtgtacatgtgcagaaatgggatggaaaaatacaacacca 375
Db 308 TAACCTTAGATGGGGGTGCTCTGTGTACATGTGCAGAAATGGGATGGGAAATCAACACCA 367
Qy 376 taaagagaaaacagagagatataaactggtggtggaatgcgtcatgaaagcgtcactt 435
Db 368 TAAAGAGAAAACGAGAGGATGATAAATCTGGTGGTGAATGCGTCATGAAGGCGTCATT 427
Qy 436 ccacgagagtttatgagagagcataagccaaggagcgttgacctggactggaagttcgcatt 495
Db 428 CCACGAGAGTTTATGAGAGAGCATTAAGCCAAGGAGCTTGACCTGGACTGAAGTTCCGAT 487
Qy 496 tgaactctacaacattctgtgggatatattgttcaaaaagatatgtgtgttttccctgat 555
Db 488 TGAACCTTACACATCTCTGGGATATATGTTCAAAAAGATATTTGTTGTTTTCATGAT 547
Qy 556 ttgacaagcaagtaatttctcccaagctgattttattcaatatgtgtacgttgg--tta 613
Db 548 TTAGCAACAACTAATTTCTCCCAAGCTGATTTTATCAATATGTTAGTTGTTGTTAA 607
Qy 614 aataacttttttagatttta 633
Db 608 ATAAACTTTTTTTAGATTTA 627

RESULT 13
LOCUS BG289969 632 bp mRNA linear EST 21-FEB-2001
DEFINITION 602381388F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499181 5',
mRNA sequence.
ACCESSION BG289969
VERSION BG289969.1 GI:13046293
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10362 row: o column: 22
High quality sequence stop: 627.
Location/Qualifiers
1. 632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4499181"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"

```

/note="Organ: bladder; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 211 a 106 c 144 g 171 t  
ORIGIN

Query Match 92.2%; Score 584.4; DB 10; Length 632;  
Best Local Similarity 98.8%; Pred. No. 5.5e-147; Indels 1; Gaps 1;  
Matches 599; Conservative 0; Mismatches 6;

Qy 22 ttccttccaccttgtagaataatccttagaataactcacaataatgtgtgatctttttag 81  
|||||  
Db 1 TTCCTTCTCACCTTGAGAGAAATCCTAGAAAACCTCAGAAAATGTGTGATCTTTGTAG 60  
Qy 82 gtacttggaacttgctccagtgaaacttgaaacttgatgattatataaagaagtagagtg 141  
|||||  
b 61 GTACTGGAAACTTGTCTCCAGTGAACCTTTGTGATGATTATATGAAAGAACTAGAGTGG 120  
Qy 142 gcttggcaccagaaagtggctggcatggcgaacacccaatcagatgcatcagtggaatg 201  
|||||  
Db 121 GCTTGGCACCAGAAAGTGGCTGGCATGGCCAAACCTAACATGATCATCAGTGTGAATG 180  
Qy 202 gggatgtgatacaccattaaatctgaaagtacaccttaaaaaactgagatttccttcatac 261  
|||||  
Db 181 GGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAAAATACTGAGATTTCTTCTATAC 240  
Qy 262 tgggcccaggaatttgacgaagctcactgcagatgacagaaagtcagagacacataacct 321  
|||||  
Db 241 TGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAAGTCAAGAGCACCATAACCT 300  
Qy 322 tagatgggggtgctggtacatgtcagaaatggatggaaatcaaccaccataaaga 381  
|||||  
Db 301 TAGATGGGGTGTCTGGTACATGTGCAGAAATGGGATGGGAAATCAACACCATTAAGA 360  
Qy 382 gaaacagagaggatgataaaactggtggaatgcgtcatgaaaggcgtcacttccacga 441  
|||||  
Db 361 GAAACAGAGAGGATGATAAACTGGTGGTGAATCGGTCAATGAAAGGCGTCACTTCCACGA 420  
Qy 442 gagtttatgagagacataaagcaagggacgttgacactggactgaagtgcagttgaact 501  
|||||  
Db 421 GAGTTTATGAGAGACATAAGCCAAAGGACGCTGACCTGGACTGAAATTCGCAATGAACT 480  
Qy 502 ctacaacttctgggatattgttc-aaaaagatatgttcttccctgattagc 560  
|||||  
b 481 CTACAACATCTGTGGGATATATTGTTCAAAAAGATATTGTCGTTTCCATGATTTAGC 540  
Qy 561 aagcaagtaatttctcccaagctgattttatttcaatatggttacgttggtgtaataaact 620  
|||||  
Db 541 AAGCAACTAATTTCTCCCAAGCTGATTTTATTCAATATGTTACGTTGGTTAAATAAAC 600  
Qy 621 tttttt 626  
|||||  
Db 601 TTTTtt 606

RESULT 14  
BE379934  
LOCUS 601159424F2 NIH\_MGC\_53 608 bp mRNA linear EST 21-JUL-2000  
DEFINITION mRNA sequence.  
ACCESSION BE379934  
VERSION BE379934.1 GI:9325299  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 608)  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLCM196 row: P column: 06  
High quality sequence stop: 597.  
Location/Qualifiers  
1. 608

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3511157"  
/clone\_lib="NIH\_MGC\_53"  
/tissue\_type="carcinoma, cell line"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: bladder; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (ggcgcctggcc); Site\_2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGGCGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."  
196 a 107 c 142 g 163 t

BASE COUNT  
ORIGIN

Query Match 92.1%; Score 584.2; DB 10; Length 608;  
Best Local Similarity 99.3%; Pred. No. 6.2e-147; Indels 1; Gaps 1;  
Matches 597; Conservative 0; Mismatches 3;

Qy 16 tgcagcttccttccacttgaaataatccttagaataactcacaataatgtgtgatgctt 75  
|||||  
Db 1 TGCAGCTTCCTTCTCACCTTGAAGAAATATCTTAGAAACTCACAATAATGTGTGATGCTT 60  
Qy 76 ttgtagttacctgaaactgtctccagtgaaactttgattgattatataaagaagtag 135  
|||||  
Db 61 TTGTAGGTACCTGGAAACTTGTCTCCAGTGAACCTTTTGATGATTATATGAAGAAGTAG 120  
Qy 136 gagtgggctttgccaccagaaagtgcgtggcatggcgaacccaactaactatcatcagtg 195  
|||||  
Db 121 GAGTGGGCTTTGCCACCAGGAAAGTGGCTGGCCAAACCTTAACATGATCATCAGTG 180  
Qy 196 tgaatggggatgtgatcaccaattaaatctgaaagtacatttcaaaactagatgttct 255  
|||||  
Db 181 TGAATGGGATGTGATCACCATTAAATCTGAAGTACCTTTAAAAATACTGAGATTTCCT 240  
Qy 256 tcacttgggccaggaaatttgacgaagtctcagatgacagaaagtcagagaaagtcagagc 315  
|||||  
Db 241 TCATACTGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGAAAGTCAAGAGCACCA 300  
Qy 316 taaccttagatgggggtgctcctgtacatgtgcagaataatggatgggaaatacaaccacca 375  
|||||  
Db 301 TAACCTTAGATGGGGGTGTCTGTGTACATGTGCAGAAATGGGATGGAATAATCAACACCA 360  
Qy 376 taagagaacacgagagatgataaactggtgggaatcgctcatgaaagcgctcactt 435  
|||||  
Db 361 TAAAGAAAACGAGAGATGATAAACTGGTGGTGGATGCGTCTCATGAAAGCGGTCACTT 420  
Qy 436 ccacgagagtttatgagagacataaagcgaaggacgttgacctggactgaagtctcgat 495  
|||||  
Db 421 CCACGAGAGTTTATGAGAGAGCATAAAGCAAGGACGCTTGACCTGGACTGAAGTTCGCAT 480  
Qy 496 tgaactctacaacattctgtgggatatattgttcaaaaaagatatgtgttttccctgat 555  
|||||  
Db 481 TGAACCTCTACAACATCTGTGGGATATATTGTTCAAAAAGATATTGTCGTTTCCATGAT 540

QY 556 ttgcaagcaagtaatttctcccaagctgattttatttcaatat-ggttacgttggttaa 614  
|||||  
Db 541 TTAGCAAGCAACTAAATTTCTCCCAAGCTGATTTTATTCAATATGGTTACGTTGGTTAA 600  
  
QY 615 a 615  
Db 601 A 601  
  
RESULT 15  
BG287532 654 bp mRNA linear EST 21-FEB-2001  
LOCUS 602384544F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4513571 5',  
DEFINITION mRNA sequence.  
ACCESSION BG287532  
VERSION BG287532.1 GI:13041458  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 654)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10400 row: g column: 12  
High quality sequence stop: 614.  
Location/Qualifiers  
1. 654  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4513571"  
/clone\_lib="NIH\_MGC\_93"  
/tissue\_type="transitional cell papilloma, cell line"  
/lab\_host="DHI0B (phage-resistant)"  
/note="Organ: bladder; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
218 a 115 c 149 g 171 t 1 others

EE COUNT  
IGIN

Query Match 91.9%; Score 582.8; DB 10; Length 654;  
Best Local Similarity 99.2%; Pred. No. 1.5e-146;  
Matches 606; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
  
QY 16 tgcagcttcctctccacttgaaagataatcctagaaaactcacaaaatgtgtgatgctt 75  
|||||  
Db 8 TGCAGCTTCCTCTCACCCTTGAGAAATAATCCCTAGAAAACCTCACAAAATGTGTGATGC-T 66  
  
QY 76 ttgtagtaacctgaaactgtctccagtgaaaactttgatgatttatataaagaagtag 135  
|||||  
Db 67 TTGTAGGTACTCGAAACTTGTCTCCAGTGAATACTTTGATGATTATATGAAAAGAGTAG 126  
  
QY 136 gagtggctttgcacagaaagtggtgctggcagcccaaacctaatcatcatcagtg 195  
|||||  
Db 127 GAGTGGGCTTTGGCCACAGGAAAGTGGCTGGCATGGCCAAACCTTAACATCATCATCAGTG 186  
  
QY 196 tgaatggggatggtgaccattaaatctgaagatccttttaaaatactgagatttcc 255  
|||||  
Db 187 TGAATGGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAATAATCTGAGATTTCCT 246

QY 256 tcatactgggcccaggaattttgacgaagtaactcactgcagatgcagaggaagtaagagcacca 315  
|||||  
Db 247 TCATACTGGGCCAGGAATTTGACGAAGTCACTGCGAGATGACAGGAAAGTCAAGAGCACCA 306  
  
QY 316 taaccttagatgggggtgctcctggtacatgtgcagaaatgggatggaaaatacaaccacca 375  
|||||  
Db 307 TAACTTTAGATGGGGGTGCTCTGGTACATGTGCAGAAATGGGATGGAAATCAACCCACCA 366  
  
QY 376 taaagagaaaacgagagagatgataaactgggtgggaatgcgtcatgaaaggcgtcactt 435  
|||||  
Db 367 TAAAGAGAAAACGAGAGAGATGATAAACTGGTGGTGGATCGCTCATGAAAGCGCTCACTT 426  
  
QY 436 ccacgagagtttatgagagagcataaagccaaaggacgttgacctggagctgaagtgcgat 495  
|||||  
Db 427 CCACGAGAGTTTATGAGAGAGCATAAGCAAGGACGTTGACCTGGACTGAAGTTCCGAT 486  
  
QY 496 tgaactctacaacattctctgggatataattgttcaaaaagatatgtgttttccctgat 555  
|||||  
Db 487 TGAACCTTACAACATTCTGTGGGATATATNGTTCAAAAAGATATGT-TGTTTTCCATGAT 545  
  
QY 556 tttagcaagcaagtaattttctcccaagctgattttatttcaatatggttacgttggttaaa 615  
|||||  
Db 546 TTAGCAAGCAACTAAATTTCTCCCAAGCTGATTTTATTCATATATGTTACGTTGGTTAAA 605  
  
QY 616 taacttttttt 626  
|||||  
Db 606 TAACTTTTTTT 616

Search completed: June 8, 2002, 09:32:01  
Job time: 9899 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 07:57:32 ; Search time 102.09 Seconds  
(without alignments)  
1525.435 Million cell updates/sec

Title: US-09-503-596-2  
Perfect score: 634  
Sequence: 1 ggaattccaggagggtgcag.....ataacttttttagatttag 634

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	155.6	24.5	408	1	US-08-446-600A-3
2	143.2	22.6	731	1	US-09-043-646-1
3	63.6	10.0	7218	1	US-08-232-463-14
4	60.8	9.6	515	4	US-08-385-982-474
5	59.6	9.4	640	2	US-08-847-724-2
6	50.2	7.9	924	1	US-08-468-709B-1
7	50.2	7.9	924	2	US-08-241-664B-1
8	50.2	7.9	924	5	PCT-US93-03936-1
9	48.8	7.7	525	1	US-08-468-709B-3
10	48.8	7.7	525	2	US-08-241-664B-3
11	48.8	7.7	525	5	PCT-US93-03936-3
12	46.4	7.3	560	4	US-09-040-984-72
13	46.4	7.3	560	4	US-09-123-912-72
14	46.4	7.3	944	2	US-08-820-825-1
15	46.4	7.3	944	4	US-09-307-817-1
16	44.8	7.1	957	3	US-08-899-031-2
17	41.8	6.6	1717	1	US-08-468-709B-6
18	41.8	6.6	1717	2	US-08-241-664B-6
19	41.8	6.6	1717	5	PCT-US93-03936-6
20	35.6	5.6	2263	4	US-08-687-865A-1
21	35.6	5.6	2263	4	US-09-043-711-1
22	33.6	5.3	289	4	US-09-007-005-17
23	33.6	5.3	289	4	US-09-244-796-17
24	32.6	5.1	530	4	US-09-328-111-682
25	32.4	5.1	613	4	US-08-998-416-205
26	32.2	5.1	593	4	US-09-328-111-724
27	32.2	5.1	650	4	US-09-328-111-789

C 28	32	5.0	4134	4	US-09-196-387-1	Sequence 1, Appli
C 29	32	5.0	4491	4	US-09-196-387-7	Sequence 7, Appli
C 30	32	5.0	14507	3	US-08-785-150-1	Sequence 1, Appli
C 31	32	5.0	14507	4	US-09-660-299-1	Sequence 1, Appli
C 32	32	5.0	14507	4	US-09-435-377-1	Sequence 1, Appli
C 33	31.8	5.0	3930	3	US-09-162-373-2	Sequence 2, Appli
C 34	31.8	5.0	3930	4	US-09-467-946-2	Sequence 2, Appli
C 35	31.4	5.0	1418	2	US-08-922-171-1	Sequence 1, Appli
C 36	31.4	5.0	3395	4	US-09-103-478-3	Sequence 3, Appli
C 37	31.4	5.0	3395	4	US-09-193-931C-3	Sequence 3, Appli
C 38	31.4	5.0	7560	4	US-09-103-478-4	Sequence 4, Appli
C 39	31.4	5.0	7560	4	US-09-193-931C-4	Sequence 4, Appli
C 40	30.8	4.9	1985	3	US-09-173-581-17	Sequence 17, Appli
C 41	30.8	4.9	1985	4	US-09-420-915-17	Sequence 17, Appli
C 42	30.8	4.9	2074	2	US-09-018-576-2	Sequence 2, Appli
C 43	30.8	4.9	2074	2	US-09-018-576-11	Sequence 11, Appli
C 44	30.8	4.9	2074	3	US-09-248-137-2	Sequence 2, Appli
C 45	30.8	4.9	2074	3	US-09-248-137-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1  
US-08-446-600A-3  
; Sequence 3, Application US/08446600A  
; Patent No. 5719126  
; GENERAL INFORMATION:  
; APPLICANT: No. 5719126dlund, James J. and Farooqui, Jamal Z.  
; TITLE OF INVENTION: MELANOGENIC INHIBITOR, AND METHODS OF PRODUCING AND USING T  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Frost & Jacobs  
; STREET: 2500 PNC Center, 201 East Fifth St.  
; CITY: Cincinnati  
; STATE: OH  
; COUNTRY: USA  
; ZIP: 45202-4182  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,600A  
; FILING DATE: 24 May 1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/980,513  
; FILING DATE: 24 No. 5719126ember 1992  
; APPLICATION NUMBER: PCT/US93/11139  
; FILING DATE: 16 No. 5719126ember 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ann G. Robinson  
; REGISTRATION NUMBER: 39,820  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (513) 651-6128  
; TELEFAX: (513) 651-6981  
; TELEX: 21-4396 P&J Gin  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 408 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Unknown  
; US-08-446-600A-3

Query Match 24.5%; Score 155.6; DB 1; Length 408;  
Best Local Similarity 62.7%; Pred. No. 1.9e-39;  
Matches 242; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

Qy 76 ttgtaggacacgaaactgtctccagtgaaaaactttgtagtatttatatgaaagaagtag 135  
Dy 20 TGAAGAGAGATGGCGCTGTGGACAGCAAGGCTTTGTATGATACATGAGGAGCTAG 79  
Qy 136 gagggtggcttgcaccacggaagtggtgcatgagcaaaccttaacatgatcatcagtg 195  
Dy 80 GAGTGGGAATAGCTTTGCGAATAATGGCGCAATGGCCCAAGCCAGATTGTATCATCACTT 139  
Qy 196 tgaatgggagtgatcaccattaaactgaaagtacattttaaataactagatttctt 255  
Dy 140 GTGATGGTAAACCTACCAATAAACTGAGAGACTTTGAAACACACAGTTTTCTT 199  
Qy 256 tcacttggcgaggaatttgcacgaagtcactgcagatgacaggaagtcgaagacacca 315  
Dy 200 GTACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAGAAAAAATCAGACTGCT 259  
Qy 316 taaacttagatgggggtgctctgtacatgtgcgaaatgggaggaatgaaacacacca 375  
Dy 260 GCAACTTTTACAGATGGTGCATTTGCTCAGCATCAGGATGGGATGGGAGGAAAGCACAA 319  
Qy 376 taagagaaaaacgagagatgataaactgtgtggaatgcgtatgaaaggcgtcactt 435  
Dy 320 TAACAAGAAATTAAGATGGGAATTAAGTGGTGGAGTGTGTCATGAACAATGTCACCT 379  
Qy 436 ccacgagagtttatgagagacataa 461  
Dy 380 GTACTCGGATCTATGAAAAAGTAGAA 405

## RESULT 2

US-09-043-646-1  
; Sequence 1, Application US/09043646  
; Patent No. 6329169  
; GENERAL INFORMATION:  
; APPLICANT: Ni, Jian  
; APPLICANT: Yu, Guo-Liang  
; APPLICANT: Gentz, Reiner L.  
; TITLE OF INVENTION: Cytostatin II  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: MD  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/043,646  
; FILING DATE: 09-Sep-98  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/12540  
; FILING DATE: 29-Sep-95  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michele M. Wales  
; REGISTRATION NUMBER: 43,975  
; REFERENCE/DOCKET NUMBER: PF221US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-309-8504  
; TELEFAX: 301-309-8439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 731 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: 16..411  
US-09-043-646-1

Query Match 22.6%; Score 143.2; DB 4; Length 731;  
Best Local Similarity 59.2%; Pred. No. 2e-35;  
Matches 244; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

Qy 50 gaaaactcacaaaatgtgtgatgctttttagtaggtaccctggaaaacttgcctccagtgaaaa 109  
Dy 3 GGAAGAGGCAAGGATGGTGGAGGCTTTCTGCTACCTGGAAAGCTGACCAACAGTCAGAA 62  
Qy 110 cttgatgattatgaaagaagtagtggtgcttccaccaggaagtggtcgtgcat 169  
Dy 63 CTTTGATGATGATACATGAAGGCTCTAGGCGTGGGCTTTGCCACTAGGAGTGGGAATGT 122  
Qy 170 ggcacaaacctaacatgatcatcagtgatgaaatggggatgtgatcaccattaaatctgaaag 229  
Dy 123 GACCAAAACCAACGGTAATTATCAGTCAAGAAAGGAGACAAAGTGGTGCACGAGACTCTCAG 182  
Qy 230 tacctttaaaaaactagagatttccctcactactggccagggaatttgacgaagtcactgc 289  
Dy 183 CACATTCAAGAACACGGAGATTAGTTCCAGCTGGGGAAGAGATTGTGATGAACCACTGC 242  
Qy 290 agatgacaggaagtcgaagagcaccataaccttagatgggggtgtcctgtgtacatgtgca 349  
Dy 243 AGATGATAGAAACTGTAAGTCTGTGTAGCTGGCTGGAGACAAACTTGTTTCACATACA 302  
Qy 350 gaaatgggagtgaaatacaaccacataaagagaaaaagagagtagatgataactgggtggt 409  
Dy 303 GAAATGGGATGGCAAGAAACAAATTTTGTAAAGAAATTAAGGATGCGCAAAATGGTTAT 362  
Qy 410 ggaatgctcatgaaaggcgtcacttccacgagagatttatgagagacataa 461  
Dy 363 GACCTTACTTTTGTGTGATGTGGTGTCTGCTGCGCACATATGAGAGAGGCATAA 414

## RESULT 3

US-08-232-463-14/C  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU



;  
; LENGTH: 640 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: THYNOT02  
; CLONE: 879706  
US-08-847-724-2

Query Match 9.4%; Score 59.6; DB 2; Length 640;  
Best Local Similarity 55.9%; Pred. No. 3.2e-09;  
Matches 113; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 81 ggtacctggaactgtctccagtgaaactttgatgattatgaaagaagtagagtg 140  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 56 GGTACTTGGACCTGCTCAGCAGCCACACACTTCGAGGCTACATGCTGCCCTAGTATT 115  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 141 ggccttgcacaggaagtgctggcagtgccaaacccaacatgatcatcagtgta 200  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 116 GACTTTGCCACTCGTAAATAGCCAAAGTTGCTGAAGCCACACAGAAAGTGATTGAGCAGAAT 175  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 201 ggggatgtgatacaccataaactgaaagtagtacctttaaaaatactgagatttcctcata 260  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 176 GGGGATCTTTTACCATCCACACAGCAGCAGCCCTAAGGAACACTACTTTGTGAATTTAAA 235  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 261 ctggggccaggaatttgacgaag 282  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 236 GTTGGAGAAGATTGTGATGAG 257  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

## RESULT 6

US-08-468-709B-1  
; Sequence 1, Application US/08468709B  
; Patent No. 5654137  
; GENERAL INFORMATION:  
; APPLICANT: Astrom, Anders  
; APPLICANT: Voorhees, John  
; APPLICANT: Pettersson, Ulrika  
; APPLICANT: Cromie, Matthew  
; APPLICANT: Tavakkol, Amir  
; TITLE OF INVENTION: HUMAN CRABP-I AND CRABP-II  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
; STREET: PO Box 828  
; CITY: Bloomfield Hills  
; STATE: Michigan  
; COUNTRY: United States of America  
; ZIP: 48303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,709B  
; FILING DATE: 06/06/95  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Deann F.  
; REGISTRATION NUMBER: 36,683  
; REFERENCE/DOCKET NUMBER: 2115-00676DVF  
; TELEPHONE: (810) 641-1600  
; TELEFAX: (810) 641-0270  
; TELEX: 287637  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 924 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA to mRNA

;  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; TISSUE TYPE: skin  
; CELL TYPE: fibroblast  
; IMMEDIATE SOURCE:  
; LIBRARY: HUMAN SKIN FIBROBLAST LAMBDA GT11  
; CLONE: LAMBDA FL1

;  
; NAME/KEY: CDS  
; LOCATION: 99..515  
; OTHER INFORMATION: /codon\_start= 99  
; OTHER INFORMATION: /citation= ({1})  
; FEATURE:  
; NAME/KEY: polyA\_site  
; LOCATION: 924  
; OTHER INFORMATION: /citation= ({1})  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: 1..98  
; OTHER INFORMATION: /citation= ({1})  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: 516..924  
; OTHER INFORMATION: /citation= ({1})  
; FEATURE:  
; NAME/KEY: polyA\_signal  
; LOCATION: 911..916  
; OTHER INFORMATION: /citation= ({1})  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: 513..515  
; OTHER INFORMATION: /citation= ({1})  
; PUBLICATION INFORMATION:  
; AUTHORS: Astrom, Anders  
; AUTHORS: Tavakkol, Amir  
; AUTHORS: Pettersson, Ulrika  
; AUTHORS: Cromie, Matthew  
; AUTHORS: Elder, James T.  
; AUTHORS: Voorhees, John J.  
; TITLE: Molecular Cloning of Two Human Cellular  
; TITLE: Retinoic Acid-Proteins (CRABP)  
; JOURNAL: J. Biol. Chem.  
; DATE: 1991

;  
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 924  
; PUBLICATION INFORMATION:  
; AUTHORS: Astrom, Anders  
; AUTHORS: Tavakkol, Amir  
; AUTHORS: Elder, James T.  
; AUTHORS: Pettersson, Ulrika  
; AUTHORS: Cromie, Matthew  
; AUTHORS: Voorhees, John J.  
; TITLE: Cloning of CRABP II cDNA from Human Skin:  
; TITLE: Retinoic Acid Induces Expression of CRABP II but  
; TITLE: No. 5654137 CRABP in Human Skin in Vivo and in Dermal but  
; TITLE: No. 5654137 Lung Fibroblasts in Vitro  
; JOURNAL: J. Invest. Dermatol.  
; VOLUME: 96  
; PAGES: 547-547  
; DATE: April-1991

US-08-468-709B-1

Query Match 7.9%; Score 50.2; DB 1; Length 924;  
Best Local Similarity 54.0%; Pred. No. 3.5e-06;  
Matches 128; Conservative 0; Mismatches 103; Indels 6; Gaps 1;

Qy 81 ggtacctggaactgtctccagtgaaactttgatgattatgaaagaagtagagtg 140  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 114 GGCAACTGGAATAATCATCCGATCGGAAACCTCGAGGAATGCTCAAAGTGTGGGGGTG 173  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Qy 141 ggccttgcacaggaa-----agtggctggcagtgccaaacccaacatgatcatcagt 194

Db 174 AATGTCATGCTGAGGAGAGATTGCTGTGGCTGCAGCGTCCAGCCACCGACCTGGAGATCAAA 233  
QY 195 gtgaatggggatgtgatccaccattaaatctgaagtagctcttttaaaataactagagattcc 254  
Db 234 CAGGAGGAGACACTTCTACATCAAAACCTCCACCACCGTGGCCACACAGAGATTAAC 293  
QY 255 ttcatctggccaggaaatttgcagagtcactgcagatgcagagaaagtcaagagc 311  
Db 294 TTCAAGGTTGGGAGGAGTTTGAGGAGCAGACTGTGGTGGGAGGCCCTGTAAGAGC 350

## RESULT 7

US-08-241-664B-1  
; Sequence 1, Application US/08241664B  
; Patent No. 5871909  
; GENERAL INFORMATION:  
; APPLICANT: Voorhees, John J.  
; APPLICANT: Astrom, Anders  
; APPLICANT: Patterson, Ulrika  
; APPLICANT: Tavakkol, Amir  
; TITLE OF INVENTION: HUMAN CRABP-I AND CRABP-II  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
; STREET: PO Box 828  
; CITY: Bloomfield Hills  
; STATE: Michigan  
; COUNTRY: United States of America  
; ZIP: 48303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,664B  
; FILING DATE: May 11, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Deann F.  
; REGISTRATION NUMBER: 36,683  
; REFERENCE/DOCKET NUMBER: 2115-00676COD  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (810) 641-1600  
; TELEFAX: (810) 641-0270  
; TELEX: 287637  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 924 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapien  
; TISSUE TYPE: skin  
; CELL TYPE: fibroblast  
; IMMEDIATE SOURCE:  
; LIBRARY: HUMAN SKIN FIBROBLAST LAMBDA GT11  
; CLONE: LAMBDA FL1.1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 99..515  
; OTHER INFORMATION: /codon\_start= 99  
; OTHER INFORMATION: /citation= ([1])  
; FEATURE:  
; NAME/KEY: polyA\_site  
; LOCATION: 924  
; OTHER INFORMATION: /citation= ([1])  
; FEATURE:

; NAME/KEY: 5'UTR  
; LOCATION: 1..98  
; OTHER INFORMATION: /citation= ([1])  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: 516..924  
; OTHER INFORMATION: /citation= ([1])  
; FEATURE:  
; NAME/KEY: polyA\_signal  
; LOCATION: 911..916  
; OTHER INFORMATION: /citation= ([1])  
; FEATURE:  
; NAME/KEY: terminator  
; LOCATION: 513..515  
; OTHER INFORMATION: /citation= ([1])  
; PUBLICATION INFORMATION:  
; AUTHORS: Astrom, Anders  
; AUTHORS: Tavakkol, Amir  
; AUTHORS: Patterson, Ulrika  
; AUTHORS: Cromie, Matthew  
; AUTHORS: Voorhees, John J.  
; TITLE: Molecular Cloning of Two Human Cellular  
; TITLE: Retinoic Acid-Proteins (CRABP)  
; JOURNAL: J. Biol. Chem.  
; DATE: 1991  
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 924  
; PUBLICATION INFORMATION:  
; AUTHORS: Astrom, Anders  
; AUTHORS: Tavakkol, Amir  
; AUTHORS: Elder, James T.  
; AUTHORS: Patterson, Ulrika  
; AUTHORS: Cromie, Matthew  
; AUTHORS: Voorhees, John J.  
; TITLE: Cloning of CRABP-II cDNA from Human Skin:  
; TITLE: Retinoic Acid Induces Expression of CRABP-II but  
; TITLE: No. 5871909 CRABP-I in Human Skin in Vivo and in Dermal but  
; TITLE: No. 5871909 Lung Fibroblasts in Vitro  
; JOURNAL: J. Invest. Dermatol.  
; VOLUME: 96  
; PAGES: 547-547  
; DATE: April-1991  
; US-08-241-664B-1

Query Match 7.9%; Score 50.2; DB 2; Length 924;  
Best Local Similarity 54.0%; Pred. No. 3.5e-06;  
Matches 128; Conservative 0; Mismatches 103; Indels 6; Gaps 1;

QY 81 ggtacctggaaacttgtctccagtgaaactttgatgattatgaagaagtaggagtg 140  
Db 114 GGCACCTGGAAAATCATCCGATCGGAAAACCTTCGAGGAATTGCTCAAAAGTGGGGGTG 173  
QY 141 ggccttgccaccaggaa-----agtggctggcatggcacaacctaaacatgatacagt 194  
Db 174 AATGTGATGCTGAGGAGAGATTGCTGTGGCTGCAGCGTCCAGCGCAGCAGCTGGAGATCAA 233  
QY 195 gtgaatggggatgtgatcaccattaaatctgaaagtacattttaaataactgagattcc 254  
Db 234 CAGGAGGAGACACTTTCTACATCAAAACCTCCACCACCGCTGCCACACACAGAGATTAAC 293  
QY 255 ttcatctggccaggaaatttgcagagtcactgcagatgcagagaaagtcaagagc 311  
Db 294 TTCAAGGTTGGGAGGAGTTTGAGGAGCAGACTGTGGTGGGAGGCCCTGTAAGAGC 350

## RESULT 8

PCT-US93-03936-1  
; Sequence 1, Application PC/TUS9303936  
; GENERAL INFORMATION:  
; APPLICANT: Voorhees, John J.  
; APPLICANT: Astrom, Anders  
; APPLICANT: Patterson, Ulrika

APPLICANT: Tavakkol, Amir  
 TITLE OF INVENTION: HUMAN CRABP-I AND CRABP-II  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Harness, Dickey & Pierce  
 STREET: PO Box 828  
 CITY: Bloomfield Hills  
 STATE: Michigan  
 COUNTRY: United States of America  
 ZIP: 48013  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US93/03936  
 FILING DATE: 19930427  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lewak, Anna M.  
 REGISTRATION NUMBER: 33,006  
 REFERENCE/DOCKET NUMBER: 211500676POB  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (313) 641-1600  
 TELEFAX: (313) 641-0270  
 TELEX: 287637  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 924 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapien  
 TISSUE TYPE: skin  
 CELL TYPE: fibroblast  
 IMMEDIATE SOURCE:  
 LIBRARY: HUMAN SKIN FIBROBLAST LAMBDA GT11  
 CLONE: LAMBDA F1.1  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 99..515  
 OTHER INFORMATION: /codon\_start= 99  
 OTHER INFORMATION: /citation= ([1])  
 FEATURE:  
 NAME/KEY: polyA\_site  
 LOCATION: 924  
 OTHER INFORMATION: /citation= ([1])  
 FEATURE:  
 NAME/KEY: 5'UTR  
 LOCATION: 1..98  
 OTHER INFORMATION: /citation= ([1])  
 FEATURE:  
 NAME/KEY: 3'UTR  
 LOCATION: 516..924  
 OTHER INFORMATION: /citation= ([1])  
 FEATURE:  
 NAME/KEY: polyA\_signal  
 LOCATION: 911..916  
 OTHER INFORMATION: /citation= ([1])  
 FEATURE:  
 NAME/KEY: terminator  
 LOCATION: 513..515  
 OTHER INFORMATION: /citation= ([1])  
 PUBLICATION INFORMATION:  
 AUTHORS: Astrom, Anders  
 AUTHORS: Tavakkol, Amir  
 AUTHORS: Pettersson, Ulrika  
 AUTHORS: Cromie, Matthew

AUTHORS: Elder, James T.  
AUTHORS: Voorhees, John J.  
TITLE: Molecular Cloning of Two Human Cellular  
TITLE: Retinoic Acid-Proteins (CRABP)  
JOURNAL: J. Biol. Chem.  
DATE: 1991  
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 924  
PUBLICATION INFORMATION:  
AUTHORS: Astrom, Anders  
AUTHORS: Tavakkol, Amir  
AUTHORS: Elder, James T.  
AUTHORS: Pettersson, Ulrika  
AUTHORS: Cromie, Matthew  
AUTHORS: Voorhees, John J.  
TITLE: Cloning of CRABP II cDNA from Human Skin:  
TITLE: Retinoic Acid Induces Expression of CRABP II but  
TITLE: Not CRABP I in Human Skin in Vivo and in Dermal but  
TITLE: Not Lung Fibroblasts in Vitro  
JOURNAL: J. Invest. Dermatol.  
VOLUME: 96  
PAGES: 547-547  
DATE: April-1991  
PCT-US93-03936-1

Query Match	7.9%	Score 50.2	DB 5	Length 924	
Best Local Similarity	54.0%	Pred. No. 3.5e-06			
Matches 128	Conservative	0	Mismatches 103	Indels 6	Gaps
QY	81	ggfacctgaaacttctccagtgaaaactttgatgattatatgaagaagtagagtg	140		
Db	114	GGCAACTGGAAATCATCCGATCGGAAACITTCGAGGAATTGCTCAAGTGTGGGGTG	173		
QY	141	ggotttgcaccaggaa-----agtggctggcagtggccaaacctatacatgatcagt	194		
Db	174	AATGTGATGCTGAGGAAGATTGCTGTGGCTGGAGGCTCCAGCCAGCATGGAGATCAA	233		
QY	195	gtgaatggggatgtagtcaccattaaactgaagtacacctttaaaaatactgagattcc	254		
Db	234	CAGGAGGGAGACACTTTCATCATCAAACTCCACCCTGGCCACCACAGGATTAAAC	293		
QY	255	tccatactgggccaggaatttgacgaagtcactgcagatgacaggaagtcaagagc	311		
Db	294	TTCAAGTTGGGAGGAGTTTCAGGAGCAGACTGTGGATGGGAGGCCCTGTGAAGAC	350		

RESULT 9  
US-08-468-709B-3  
Sequence 3, Application US/08468709B  
Patent No. 5654137  
GENERAL INFORMATION:  
APPLICANT: Astrom, Anders  
APPLICANT: Voorhees, John  
APPLICANT: Patterson, Ulrika  
APPLICANT: Tavakkol, Amir  
TITLE OF INVENTION: HUMAN CRABP-I AND CRABP-II  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
STREET: PO Box 828  
CITY: Bloomfield Hills  
STATE: Michigan  
COUNTRY: United States of America  
ZIP: 48303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468-709B  
FILING DATE: 06/06/95  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Deann F.  
REGISTRATION NUMBER: 36,683  
REFERENCE/DOCKET NUMBER: 2115-00676DVF  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (810) 641-1600  
TELEFAX: (810) 641-0270  
TELEX: 287637  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 525 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapien  
TISSUE TYPE: skin  
CELL TYPE: fibroblast  
IMMEDIATE SOURCE:  
LIBRARY: human skin Lambda ZapII  
CLONE: lambda s3.1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 8..418  
OTHER INFORMATION: /codon\_start= 8  
OTHER INFORMATION: /citation= ([1])  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..7  
OTHER INFORMATION: /citation= ([1])  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 419..525  
OTHER INFORMATION: /citation= ([1])  
FEATURE:  
NAME/KEY: terminator  
LOCATION: 419..421  
OTHER INFORMATION: /citation= ([1])  
PUBLICATION INFORMATION:  
AUTHORS: Astrom, Anders  
AUTHORS: Tavakkol, Amir  
AUTHORS: Pettersson, Ulrika  
AUTHORS: Cromie, Matthew  
AUTHORS: Elder, James T.  
AUTHORS: Voorhees, John J.  
TITLE: Molecular Cloning of Two Human Cellular  
TITLE: Retinoic Acid-Proteins (CRABP)  
JOURNAL: J. Biol. Chem.  
DATE: 1991  
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 525  
PUBLICATION INFORMATION:  
AUTHORS: Astrom, Anders  
AUTHORS: Tavakkol, Amir  
AUTHORS: Elder, James T.  
AUTHORS: Pettersson, Ulrika  
AUTHORS: Cromie, Matthew  
AUTHORS: Voorhees, John J.  
TITLE: Cloning of CRABP II cdna from Human Skin:  
TITLE: Retinoic Acid Induces Expression of CRABP II but  
TITLE: No. 5654137 CRABP in Human Skin in Vivo and in Dermal but  
TITLE: No. 5654137 Lung Fibroblasts in Vitro  
JOURNAL: J. Invest. Dermatol.  
VOLUME: 96  
PAGES: 547-547  
DATE: April-1991  
US-08-468-709B-3

Query Match 7.7%; Score 48.8; DB 1; Length 525;  
Best Local Similarity 53.3%; Pred. No. 7e-06;

Matches 129; Conservative 0; Mismatches 107; Indels 6; Gaps 1;  
Qy 75 ttgttagtagtacctggaactgtgtccagtgaaactttgatgattatgaagaagta 134  
Db 17 TTGCGCGGCACTGGAGATGCGCAGCAGCAGGAGAAATTCGACGAGCTGCTCAAGGCACTG 76  
Qy 135 ggagtgggtttgccaccaggaaagtggctggcatgg-----ccaaacctaacatgac 188  
Db 77 GGTGTGAACGCATGCTGAGGAAGGTGGCGTAGCGGTGGCTCCAAAGCCGACCTGGAG 136  
Qy 189 atcaagtgtgaatggggtgtgataccattataatctgaagtaccttttaaaatactag 248  
Db 137 ATCCGCCAGGCGGGATCAGTTCTACATCAAGACATCCACCACGTCGCCACCACTGAG 196  
Qy 249 attctctcactatggcaggaaatttgacgaagtcaactgcagatgacaggaaatcaag 308  
Db 197 ATCAACTTCAAGTTCGGAGAAGGCTTTGAGGAGGAGACCGTGGACGCGCAAGTGCAGG 256  
Qy 309 ag 310  
Db 257 AG 258  
RESULT 10  
US-08-241-664B-3  
; Sequence.3, Application US/08241664B  
; Patent No. 5871909  
; GENERAL INFORMATION:  
; APPLICANT: Voorhees, John J.  
; APPLICANT: Astrom, Anders  
; APPLICANT: Pettersson, Ulrika  
; APPLICANT: Tavakkol, Amir  
; TITLE OF INVENTION: HUMAN CRABP-I AND CRABP-II  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.  
; STREET: PO Box 828  
; CITY: Bloomfield Hills  
; STATE: Michigan  
; COUNTRY: United States of America  
; ZIP: 48303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,664B  
; FILING DATE: May 11, 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Deann F.  
; REGISTRATION NUMBER: 36,683  
; REFERENCE/DOCKET NUMBER: 2115-00676COD  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (810) 641-1600  
; TELEFAX: (810) 641-0270  
; TELEX: 287637  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 525 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo Sapien  
; TISSUE TYPE: skin  
; CELL TYPE: fibroblast  
; IMMEDIATE SOURCE:  
; LIBRARY: human skin Lambda ZapII





Query Match 7.3%; Score 46.4; DB 4; Length 560;  
Best Local Similarity 50.6%; Pred. No. 4.1e-05;  
Matches 120. Conservative 0. Mismatches 111. Indels

	14	ggtacatggaaactctctccagtgaaaactttgagattgatataagaagaagtagagcgc	197
DQ			
Db	138	GGCAACTGGAAAAATCATCCGACCGAAAACTTCGANGAATTGCTCNAANTGCTGGGGGTG	197
DQ	141	ggotttgcaccaggaaaa-----gtggctggcatggccaaacctaacatgatcatcgat	194
Db	198	AATGTGATGCTNANGAANATTGCTGTGGCTGCAGCGTCCAAGCCAGCAGTGAGGATCNAA	257
DQ	195	gtgaatggggatdgtatcaccaattaactgtgaagtacactttaataaatactgagattccc	254
Db	258	CAGGAGGAGACACTTTCTACATCAAACCCTCCACCACCGTGGCGCACCAAGAATTAAC	317
DQ	255	tteactactggccagaatttcagcaagtcactgcsgatgcacaggaaagtccaagagc	311
Db	318	TTCNNNGTTGGGCGAGANTTTGAGGANCAAACTGTGGATNGGANGCCTGTNAAAACC	374

```

RESULT 13
US-09-123-912-72
; Sequence 72, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123.912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (83)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (173)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (183)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (186)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base

```

LOCATION: (209)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (211)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (215)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (255)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (321)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (322)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (323)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (335)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (344)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (357)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (361)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (368)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (394)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (412)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (415)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (442)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (455)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (469)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (472)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (475)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (487)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (513)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (522)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (528)	OTHER INFORMATION: Where n is a, c, g or t
NAME/KEY: modified_base	
LOCATION: (531)	OTHER INFORMATION: Where n is a, c, g or t



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 07:59:27 ; Search time 369.19 Seconds  
(without alignments)  
2948.411 Million cell updates/sec

Title: US-09-503-596-2  
Perfect score: 634  
Sequence: 1 ggaattccaggagggtgcag.....ataacttttttagatttag 634

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1736436 seqs, 858457221 residues

al number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N:Geneseq\_032802.\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	634	100.0	634	21 AAA37717	Human AFABP cDNA s
2	594.6	93.8	709	21 AAF21643	Human breast and o
3	432.6	68.2	463	21 AAC10273	Human secreted pro
4	386.2	60.9	614	21 AAA37716	Mouse AFABP cDNA s
5	376.8	59.4	604	22 AAH34528	Human colon cancer
6	360	56.8	435	16 AAH21621	Human gene signatu
7	341.4	53.8	399	20 AAH32483	Porcine adipocyte
8	247.6	39.1	351	21 AAA37715	Mouse AFABP cDNA f
9	211.4	33.3	411	19 AAV10405	Human myelin P2 pr

10	205	32.3	276	21	AAZ42917	Human 5' EST isola
11	203	32.0	374	21	AAK86453	Mouse Ng-119K123 g
c 12	201	31.7	366	22	AAK53474	Murine transport a
13	194	30.6	493	21	AAC03853	Human secreted pro
14	159.4	25.1	8144	20	AAK32482	Porcine adipocyte
15	157.8	24.9	313	21	AAK86452	Mouse adipocyte li
16	155.8	24.6	662	22	AAK13247	Human DNA encoding
17	155.8	24.6	720	20	AAZ77538	Human ovarian tumo
18	155.8	24.6	1072	24	AAK94888	Human DNA sequence
19	155.6	24.5	408	15	AAQ66842	Melanogenic inhibi
20	143.2	22.6	613	19	AAV68587	Nucleotide sequenc
21	143.2	22.6	731	18	AAV74751	Human cytosstatin I
22	143.2	22.6	731	19	AAV68225	Nucleotide sequenc
23	143.2	22.6	754	20	AAV73483	Human 14-16 kDa FA
24	141.6	22.3	396	20	AAV73482	Human 14-16 kDa FA
25	139.2	22.0	606	21	AAC03963	Human secreted pro
26	130.2	20.5	413	22	AAK13344	Human cDNA encodin
27	130.2	20.5	499	22	AAK13343	Human cDNA encodin
28	114.2	18.0	698	23	AAK91764	DNA encoding novel
29	113	17.8	406	23	AAK15728	DNA encoding cytos
30	113	17.8	418	23	AAK15729	DNA encoding cytos
c 31	103	16.2	488	22	AAK54160	Murine transport a
32	101.4	16.0	5042	18	AAK89493	Porcine heart-fatt
c 33	98.8	15.6	343	22	AAK53737	Murine transport a
34	87.6	13.8	286	21	AAC01411	Human secreted pro
35	87	13.7	549	21	AAC93830	Cat flea hindgut a
36	75.8	12.0	687	23	ABL11747	Drosophila melanog
37	71.4	11.3	550	23	AAK15722	DNA encoding fatty
c 38	70.8	11.2	235	22	AAK53580	Murine transport a
39	70	11.0	1279	22	AAK13246	Mouse DNA encoding
40	69.4	10.9	7608	16	AAQ74118	DNA encoding mouse
41	69.4	10.9	7608	24	AB199750	Mouse ischaemic co
42	67.6	10.7	354	20	AAZ24640	Human lung tumor a
43	67.6	10.7	354	21	AAK5879	Human lung cancer-
44	63.4	10.0	2633	24	AB199378	Mouse ischaemic co
45	61.2	9.7	2392	23	AAK90327	DNA encoding novel

## ALIGNMENTS

RESULT 1  
AAA37717  
ID AAA37717 standard; cDNA; 634 BP.  
XX  
AC AAA37717;  
XX  
DT 22-NOV-2000 (first entry)  
XX  
DE Human AFABP cDNA sequence.  
XX  
KW AFABP; adipocyte fatty acid binding protein; atherosclerotic lesion;  
KW lesion formation inhibition; macrophage; adipocyte; atherosclerosis;  
KW serum cholesterol; therapy; human; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 63..461  
FT FT /\*tag= a  
FT FT /product= AFABP  
XX  
PN WO200047734-A1.  
XX  
PD 17-AUG-2000.  
XX  
PF 11-FEB-2000; 2000WO-US03560.  
XX  
PR 12-FEB-1999; 99US-0119880.  
XX (HARD ) HARVARD COLLEGE.  
XX Lee M, Perrella MA, Hotamisligil GS;

XX WPI; 2000-506094/45.  
DR P-PSDB; AAY90320.  
XX  
PT Reducing expression of adipocyte fatty acid binding protein through  
PT administration of a compound is used to inhibit formation of an  
PT atherosclerotic lesion -  
XX  
PS Disclosure; Page 10; 43pp; English.  
XX  
CC This sequence represents the human AFABP (adipocyte fatty  
CC acid binding protein) cDNA sequence. The invention relates to a method  
CC for inhibiting formation of an atherosclerotic lesion comprising  
CC administering to a mammal a compound that reduces expression of adipocyte  
CC fatty acid binding protein (AFABP). The method is used to inhibit  
CC formation of atherosclerotic lesions. The method is used to identify  
CC compounds which can be used to inhibit formation of atherosclerotic  
CC lesions through inhibition of AFABP binding to an intracellular ligand  
CC in a macrophage or adipocyte, inhibition of development of an  
CC atherosclerotic lesion, inhibition of a macrophage differentiating into  
CC a foam cell or inhibition of AFABP expression in a cell. AFABP activity  
CC may be inhibited to treat atherosclerosis or to treat individuals at risk  
CC of developing atherosclerosis. Inhibiting AFABP expression or activity  
CC reduces the development of atherosclerotic lesions despite a high level  
CC of serum cholesterol.  
XX  
SQ Sequence 634 BP; 200 A; 108 C; 151 G; 175 T; 0 other;

Query Match 100.0%; Score 634; DB 21; Length 634;  
Best Local Similarity 100.0%; Pred. No. 1.1e-179;  
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ggaattccaggaggtgcagcttctctcacttgaagaataatcctagaaaactca 60  
Db 1 ggaattccaggaggtgcagcttctctcacttgaagaataatcctagaaaactca 60  
Qy 61 aaatgtgtgctttgttagtaccctggaaactgtctccagtgaaactttgatgatt 120  
Db 61 aaatgtgtgctttgttagtaccctggaaactgtctccagtgaaactttgatgatt 120  
Qy 121 atattgaagaagtaggaggtggtcctccaccaggaagtgctgacgacaaactca 180  
Db 121 atattgaagaagtaggaggtggtcctccaccaggaagtgctgacgacaaactca 180  
Qy 181 acatgatcatcagtgtaagtggtggtgacacattaaatcgaagtgaccccttaaa 240  
Db 181 acatgatcatcagtgtaagtggtggtgacacattaaatcgaagtgaccccttaaa 240  
Qy 241 atactgagattccttatactggtggtggtggtggtggtggtggtggtggtggtg 300  
Db 241 atactgagattccttatactggtggtggtggtggtggtggtggtggtggtggtg 300  
Qy 301 aagtcagagacacataacacacacacacacacacacacacacacacacacacac 360  
Db 301 aagtcagagacacataacacacacacacacacacacacacacacacacacacac 360  
Qy 361 gaaatcaaccac 420  
Db 361 gaaatcaaccac 420  
Qy 421 tgaagggcgtcacttccacagagatttatgagagagacataagcgaagggcgtgac 480  
Db 421 tgaagggcgtcacttccacagagatttatgagagagacataagcgaagggcgtgac 480  
Qy 481 gactgaagttgcagttgaactctcaacacacacacacacacacacacacacacac 540  
Db 481 gactgaagttgcagttgaactctcaacacacacacacacacacacacacacacac 540  
Qy 541 gttgttttccctgatttagcaagcaagtaattttctccacagctgattttattcaatg 600  
Db 541 gttgttttccctgatttagcaagcaagtaattttctccacagctgattttattcaatg 600

Qy 601 gttacgttggttaataaacttttttttagatttag 634  
Db 601 gttacgttggttaataaacttttttttagatttag 634  
RESULT 2  
AAF21643  
ID AAF21643 standard; DNA; 709 BP.  
AC AAF21643;  
XX  
XX 27-MAR-2001 (first entry)  
DE Human breast and ovarian cancer associated antigen gene SEQ ID 30.  
KW Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;  
KW neoplastic; neuroprotective; antiviral; antitumor; hepatotropic;  
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;  
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;  
KW Addison's disease; allergy; autoimmune haemolytic anaemia;  
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;  
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;  
KW cardiovascular disorder; wound healing; neurological disease; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200055173-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05881.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-611515/58.  
XX  
XX P-PSDB; AAB58740.  
XX  
XX New human breast and ovarian cancer associated gene sequences and the  
XX polypeptides encoded by these genes, useful in the prevention,  
XX treatment and diagnosis of cancer, immune disorders, cardiovascular  
XX disorders and neurological diseases -  
XX  
XX Claim 1; Page 504; 1299pp; English.  
XX  
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human  
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are  
XX associated with breast and ovarian cancer. Included in the invention are  
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the  
XX isolation and characterisation of the DNA and protein sequences of the  
XX invention. The breast and ovarian cancer associated DNA, protein, agonist  
XX or antagonist sequences exhibit cytostatic; immunosuppressive;  
XX neoplastic; neuroprotective; antiviral; antitumor; hepatotropic;  
XX antidiabetic; antiinflammatory; antiparasitic and cardiant activity. The  
XX polynucleotide and protein sequences are used in the diagnosis of cancer,  
XX particularly breast and ovarian cancer. The nucleic acid sequences,  
XX proteins, agonists and antagonists may also be used in the diagnosis,  
XX prevention and treatment of immune disorders e.g. Addison's disease,  
XX allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
XX diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
XX arthritis and ulcerative colitis; cardiovascular disorders such as  
XX myocardial ischaemia; wound healing; neurological diseases such as  
XX cerebral anoxia and epilepsy; and infectious diseases.  
XX  
XX Sequence 709 BP; 227 A; 130 C; 167 G; 182 T; 3 other;

Query Match 93.8%; Score 594.6; DB 21; Length 709;  
Best Local Similarity 98.4%; Pred. No. 7.2e-168;

		Matches	611;	Conservative	0;	Mismatches	9;	Indels	1;	Gaps	1;
Qy	6	tccaggaggtgagcttctctcacttgaaataatcctgagaaacacacaaatg	65								
Db	24	tccgaaactgcagcttctctcacttgaaataatcctgagaaacacacaaatg	83								
Qy	66	tgtgatgtttgttagtgacatggaacttctcctcagtgaaacttttgattatg	125								
Db	84	tgtgatgtttgttagtgacatggaacttctcctcagtgaaacttttgattatg	143								
Qy	126	aaagaagtaggagtggtttgcccaccaggaaagtggctgcatggccaaacctaaacatg	185								
Db	144	aaagaagtaggagtggtttgcccaccaggaaagtggctgcatggccaaacctaaacatg	203								
Qy	186	atcatcagtgaaatggggatgtgatcaccattaaatcgaagtacattttaaataact	245								
Db	204	atcatcagtgaaatggggatgtgatcaccattaaatcgaagtacattttaaataact	263								
Qy	246	gagatttcttcatactgggcaggaaatttgacgaagtcactgcagatgacaggaaagtc	305								
Db	264	gagatttcttcatactgggcaggaaatttgacgaagtcactgcagatgacaggaaagtc	322								
Qy	306	aagagaccataaaccttagatgggggtgtcctgtacatgtgcagaaatgggatgaaaa	365								
Db	323	aagagaccataaaccttagatgggggtgtcctgtacatgtgcagaaatgggatgaaaa	382								
Qy	366	tcaaccaccataaagagaaaaacgagagatgataaaactggtgtggaatgcgtcatgaa	425								
Db	383	tcaaccaccataaagagaaaaacgagagatgataaaactggtgtggaatgcgtcatgaa	442								
Qy	426	ggcgtcactccacgagatttatgagagacataaagccaaggagcgttgacctggactg	485								
Db	443	ggcgtcactccacgagatttatgagagacataaagccaaggagcgttgacctggactg	502								
Qy	486	aagttcgcatgaaacttcacaaacttctgtgggatattgttcaaaaagatatgttgt	545								
Db	503	aagttcgcatgaaacttcacaaacttctgtgggatattgttcaaaaagatatgttgt	562								
Qy	546	tttccctgatttagcaagaactatttctcccaagctgattttattcaaatatgtgttac	605								
Db	563	tttccctgatttagcaagaactatttctcccaagctgattttattcaaatatgtgttac	622								
Qy	606	gttggttaataaacctttttt	626								
Db	623	gttggttaataaacctttttt	643								

RESULT 3  
AAC10273  
AAC10273 standard; cDNA; 463 BP.  
AC AAC10273;  
DT 06-OCT-2000 (first entry)  
XX Human secreted protein 5' EST, SEQ ID NO: 14348.  
DE  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX Homo sapiens.  
XX  
XX EP1033401-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 21-FEB-2000; 2000EP-0200610.  
XX  
XX 26-FEB-1999; 99US-0122487.  
XX (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.	
DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for	
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for	
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -	
XX Claim 1; SEQ ID 14348; 71pp + CD-ROM; English.	
PS The present sequence is one of a large number of 5' ESTs derived from	
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively	
CC identified within the present sequence. The 5' ESTs were prepared from	
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST	
CC sequences usually correspond mainly to the 3' untranslated region (UTR)	
CC of the mRNA because they are often obtained from oligo-dT primed cDNA	
CC libraries. Such ESTs are not well suited for isolating cDNA sequences	
CC derived from the 5' ends of mRNAs and even in those cases where longer	
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.	
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be	
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used	
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.	
CC They are used to obtain upstream regulatory sequences and to design	
CC expression and secretion vectors.	
XX Sequence 463 BP; 154 A; 87 C; 113 G; 108 T; 1 other;	
SQ	

Query Match	68.2%;	Score 432.6;	DB 21;	Length 463;
Best Local Similarity	98.5%;	Pred. No. 1.9e-119;		
Matches	446;	Conservative	1;	Mismatches 5; Indels 1; Gaps 1;
Qy	6	tccaggaggtgagcttctctcacttgaaataatcctgagaaacacacaaatg	65	
Db	12	tctgaaactgcagcttctctcacttgaaataatcctgagaaacacacaaatg	71	
Qy	66	tgtgatgtttgttagtgacatggaacttctcctcagtgaaactttgatattatg	125	
Db	72	tgtgatgtttgttagtgacatggaacttctcctcagtgaaactttgatattatg	131	
Qy	126	aaagaagtaggagtggtttgcccaccaggaaagtggctggcatggccaaacctaaatg	185	
Db	132	aaagaagtaggagtggtttgcccaccaggaaagtggctggcatggccaaacctaaatg	191	
Qy	186	atcatcagtgaaatggggatgtgatcaccattaaatcgaagtacattttaaataact	245	
Db	192	atcatcagtgaaatggggatgtgatcaccattaaatcgaagtacattttaaataact	251	
Qy	246	gagatttcttcatactgggcaggaaatttgacgaagtcactgcagatgacaggaaagtc	305	
Db	252	gagatttcttcatactgggcaggaaatttgacgaagtcactgcagatgacaggaaagtc	310	
Qy	306	aagagcaccataaaccttagatgggggtgtcctgggtacatgtgcagaaatgggatgaaaa	365	
Db	311	aagagcaccataaaccttagatgggggtgtcctgggtacatgtgcagaaatgggatgaaaa	370	
Qy	366	tcaaccaccataaagagaaaaacgagagatgataaaactggtgtggaatgcgtcatgaaa	425	
Db	371	tcaaccaccataaagagaaaaacgagagatgataaaactggtgtggaatgcgtcatgaaa	430	
Qy	426	ggcgtcactccacgagattttatgagagacga	458	
Db	431	ggcgtcactccacgagattttatgagagacga	463	

RESULT 4  
AAA37716  
ID AAA37716 standard; cDNA; 614 BP.  
XX  
XX AAA37716;  
XX  
XX 22-NOV-2000 (first entry)  
XX  
XX Mouse AFABP cDNA sequence.







XX 22-JUN-1999 (first entry)  
XX Porcine adipocyte fatty acid binding protein (A-FABP) coding sequence.  
XX A-FABP; adipocyte fatty acid binding protein; pig; porcine; allele;  
KW breeding program; fatty acid transport; fat deposition; ss.  
XX  
OS Sus scrofa.  
XX WO9914365-A1.  
PN 25-MAR-1999.  
XX  
XX 18-SEP-1998; 98WO-NL00541.  
XX 18-SEP-1997; 97EP-0202857.  
XX (DALL-) DALLAND BV.  
PA (NNEV-) NOORD NEDERLANDS VARKENSSTAMBOEK BV.  
PA (PROV-) PROVA BV.  
PA (STAM-) STAMBOEK ZUID BV.  
PA (DIER-) STICHTING INST DIERHOUDERIJ EN DIERGEZON.  
XX Gerbens F;  
PI  
XX WPI; 1999-229550/19.  
DR  
XX  
XX Nucleic acid specific for the pig adipocyte fatty acid binding  
PT protein  
XX  
XX Disclosure; Fig 2; 36pp; English.  
XX This represents the coding sequence of a pig A-FABP (adipocyte fatty acid  
CC binding protein). Nucleic acid specific for the A-FABP gene or its  
CC fragments are used (1) to localize, identify or mark alleles in the gene,  
CC particularly those associated with production traits in pigs; and (2)  
CC for specific amplification of fragments, its alleles or trait loci,  
CC particularly for distinguishing between alleles. Pigs carrying particular  
CC alleles are selected for use in breeding programs. Since A-FABP is  
CC involved in fatty acid transport and fat deposition, and may control  
CC insulin dependency, selection of particular alleles may allow regulation  
CC of fat deposition in muscle; body weight, weight gain and feed  
CC efficiency; fatty acid content of meat; embryo survival; birth rate and  
CC litter size, and milk quality and quantity.  
XX  
XX Sequence 399 BP; 133 A; 67 C; 107 G; 92 T; 0 other;

Query Match 53.88; Score 341.4; DB 20; Length 399;  
Best Local Similarity 91.08; Pred. No. 3.6e-92;  
Matches 363; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 63 atgtgtgatgtttttagtgtagctggaacattgtctccagtgaaactttgatgattat 122  
DB 1 atgtgtgatgtatttgtagtgtagctggaacattgtctccagtgaaactttgatgattac 60  
QY 123 atgaagaagtagagtggtggtttggtccaccaggaaagtggctggcgtgcaaaccttaac 182  
DB 61 atgaagaagtggtggtggtttggtccaccaggaaagtggctggcgtgcaaaccttaac 120  
QY 183 atgatcatcagtggaatgggagtgatgataccattaaatctgaaagtacatttaaaat 242  
DB 121 ctgatcatcactgtgaatgggagtgatgataccattaaatctgaaagtacatttaaaat 180  
QY 243 actgagatttcttcactggtggccaggaatttgacgaagtcactgcagatgacagaaa 302  
DB 181 actgagatgaccttcaaatgggcccaggaatttgatgaagtcactgcagatgacagaaa 240  
QY 303 gtcaagagcaccataaacttagatgggggtgtctcgtatcgtgacagaaatgggatgga 362  
DB 241 gtcaagagcaccataaacttagatgggagcgccctgtgtacaggtgacagaaatgggatgga 300

QY 363 aatcaaccacataaagagaaacagagagtagataaactggtggtggaatgcgtcatg 422  
DB 301 aagacaaccacataaagagaaacagagagtagataaactggtggtggaatgcgtcatg 360  
QY 423 aaaggcgtcacttcacagagaggttttatgagagagcataa 461  
DB 361 aaaggtgtcagcgtcaccagaatttatgagagagcataa 399

## RESULT 8

AAA37715  
ID AAA37715 standard; cDNA; 351 BP.

AC AAA37715;

DT 22-NOV-2000 (first entry)

DE Mouse AFABP cDNA fragment.

KW AFABP; adipocyte fatty acid binding protein; atherosclerotic lesion;  
KW lesion formation inhibition; macrophage; adipocyte; atherosclerosis;  
KW serum cholesterol; therapy; mouse; ss.

OS Mus sp.

PN WO200047734-A1.

PD 17-AUG-2000.

PF 11-FEB-2000; 2000WO-US03560.

PR 12-FEB-1999; 99US-0119880.

XX (HARD ) HARVARD COLLEGE.

XX Lee M, Perrella MA, Hotamisligil GS;

XX WPI; 2000-506094/45.

XX Reducing expression of adipocyte fatty acid binding protein through  
PT administration of a compound is used to inhibit formation of an  
XX atherosclerotic lesion -

PS Disclosure; Page 7; 43pp; English.

XX This sequence represents a fragment of the mouse AFABP (adipocyte fatty  
XX acid binding protein) cDNA sequence. The invention relates to a method  
XX for inhibiting formation of an atherosclerotic lesion comprising  
XX administering to a mammal a compound that reduces expression of adipocyte  
XX fatty acid binding protein (AFABP). The method is used to inhibit  
XX formation of atherosclerotic lesions. The method is used to identify  
XX compounds which can be used to inhibit formation of atherosclerotic  
XX lesions through inhibition of AFABP binding to an intracellular ligand  
XX in a macrophage or adipocyte, inhibition of development of an  
XX atherosclerotic lesion, inhibition of a macrophage differentiating into  
XX a foam cell or inhibition of AFABP expression in a cell. AFABP activity  
XX may be inhibited to treat atherosclerosis or to treat individuals at risk  
XX of developing atherosclerosis. Inhibiting AFABP expression or activity  
XX reduces the development of atherosclerotic lesions despite a high level  
XX of serum cholesterol.

XX Sequence 351 BP; 100 A; 75 C; 100 G; 76 T; 0 other;

Query Match 39.1%; Score 247.6; DB 21; Length 351;

Best Local Similarity 81.7%; Pred. No. 4.2e-64;

Matches 286; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 21 ctctctctcaccttgagaataatcctagaaaaactcacaaaaatgtgtgatttga 80

DB 2 ctctctcaccttgagaagacagctcctcctcgaagtttacaataatgtgtgatttga 61

QY 81 ggtacctggaaactgtctccagtgaaactttgatgattatgaaagagtgagt 140



sequences, corresponding to human secreted proteins. AAY64651 to AAY65438 represent the EST-related proteins corresponding to AA242265 to AA243052. The 5' ESTs can be used for producing secreted human gene products. They can be used to identify and isolate 5' untranslated regions (UTRs) and upstream regulatory regions which control the location, development stage, rate, and quantity of protein synthesis, as well as stability of mRNA. The ESTs are also useful as probes for chromosome mapping, and to obtain full length cDNA clones. The ESTs can also be used in forensic procedures to identify individuals, or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal gene expression. The products may also be used in gene therapy protocols. The nucleic acids encoding signal peptides can be used for directing extracellular secretion of a polypeptide or the insertion of a polypeptide into a membrane, or importing a polypeptide into a cell. The proteins encoded by the EST sequences may be useful in treating a variety of human conditions. Secreted proteins have therapeutic value, and the identification of new secreted proteins is valuable. AA242249 to AA242264 and AAY64644 to AAY64650 represent sequences used in the exemplification of the present invention.

Sequence 276 BP; 77 A; 71 C; 59 G; 69 T; 0 other;

Query Match 32.3%; Score 205; DB 21; Length 276;

Best Local Similarity 97.7%; Pred. No. 2.1e-51;

Matches 208; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 tccaggagggtgcagcttccttcaccttgaaagataatcctagaaaaactcacaaaaatg 65

Db 13 tctgaaactgcagcttccttcaccttgaaagataatcctagaaaaactcacaaaaatg 72

Qy 66 tctgatgtctttttaggtacctggaactgtctccagtgaaactttgatgattatg 125

Db 73 tctgatgtctttttaggtacctggaactgtctccagtgaaactttgatgattatg 132

Qy 126 aaagagtaggagtggttttgcaccaggaagtggtgcatgccaaaccttaacatg 185

Db 133 aaagagtaggagtggttttgcaccaggaagtggtgcatgccaaaccttaacatg 192

Qy 186 atcatcagtgtaagggtggtggtggtatgcaccatt 218

Db 193 atcatcagtgtaagggtggtggtggtatgcaccatt 225

RESULT 11

AAK86453

ID AAK86453 standard; cDNA; 374 BP.

XX AAK86453;

XX 01-MAR-2001 (first entry)

XX Mouse Ng-119K123 gene.

XX Growth hormone; brown adipose tissue; ds.

XX Mus sp.

XX WO200066784-A2.

XX 09-NOV-2000.

XX 05-MAY-2000; 2000WO-US12145.

XX 05-MAY-1999; 99US-0132670.

XX (UYOH-) UNIV OHIO.

XX Kopchick JJ, Li Y;

XX WPI; 2000-665338/64.

XX Diagnosis of abnormal levels of growth hormone in brown adipose tissue

PT is carried out by monitoring the transcriptional activity of one or  
XX more genes related to growth hormone activity -  
XX Disclosure; Page 59; 89pp; English.

XX The present invention provides a method of diagnosing abnormal levels of  
XX growth hormone (GH) in brown adipose tissue, by assaying the levels of  
XX specific gene transcripts. The genes of interest are those encoding  
XX glucosylphosphate isomerase, neuroleukin, pyruvate kinase, haem oxygenase,  
XX ubiquitin/ribosomal fusion protein, alpha-enolase, proteasome theta  
XX chain, trans-Golgi network protein, medium chain acyl-CoA dehydrogenase,  
XX adipocyte lipid binding protein, mitochondrial cytochrome C oxidase,  
XX mitochondrial NADH-ubiquinone oxidoreductase, mitochondrial cytochrome B  
XX or any of the genes Ng-119K123, Ng-119K15, Ng-119K36, Ng-119K62,  
XX Ng-119K42, Ng-119K58, Ng-119K65 or Ng-119K66. This is useful for  
XX diagnosing abnormal levels of GH or predicting changes in brown adipose  
XX tissue.

XX Sequence 374 BP; 130 A; 53 C; 93 G; 98 T; 0 other;

Query Match 32.0%; Score 203; DB 21; Length 374;

Best Local Similarity 78.3%; Pred. No. 9.7e-51;

Matches 281; Conservative 0; Mismatches 75; Indels 3; Gaps 3;

Qy 270 gaatttgacgaagtccactgcagatgcaggaagtcagagcaccataacaccttagatggg 329

Db 1 gaatttcgatgaataccacgcagacaggaagtggaagcaccataacaccttagatggc 60

Qy 330 ggtgtcctcgttacatgtgcagaaatgggatgaaatacaaccacataaagagaaaacga 389

Db 61 ggggcccctggtgcaggtgcagaaatgggatgaaatacaaccacataaagagaaaacga 120

Qy 390 gagatgataaactggtgtggaatgcctcatgaaagcgcacatccacagagatttat 449

Db 121 gatggtgcagagctggtgtggaatgcttatgaaagcgcacatccacagagatttat 180

Qy 450 gagagagacataagccaaagggcgttgacctggaactggaactgcgattgaaactctacaaca 509

Db 181 gaaagggcagatgagccaaaggaag-aggcctggatggaatttgcatcaaacactacaata 239

Qy 510 tctgtggggatatattgttcaaaaagata-ttgttgttttccctgatttagcaagcaagt 568

Db 240 gtcagtcggattattgttttttttaagatatgatttccactaataaagcaagcaatt 299

Qy 569 aattttctcccaagctg-atitttattcaatgtgttacgttggttaataaactttttt 626

Db 300 aatttttctgaagatgcattttattgtgatattgttattgttgatttaataaactttt 358

RESULT 12

AAK53474/c

ID AAK53474 standard; cDNA; 366 BP.

XX AAK53474;

XX 16-NOV-2001 (first entry)

XX Murine transport and binding associated protein encoding cDNA SEQ ID 39.

XX Murine; liver; gene library; amino acid synthesis; binding protein;  
XX cell metabolism; energy metabolism; fatty acid metabolism; synthesis;  
XX phospholipid metabolism; purine; pyrimidine; nucleoside; nucleotide;  
XX replication; transcription; translation; transport protein; ss.

XX Mus musculus.

XX DE20103510-UI.

XX 07-JUN-2001.

XX 28-FEB-2001; 2001DE-2003510.

```
PR 02-DEC-1999; 99DE-1058160.
XX (LION-) LION BIOSCIENCE AG.
XX WPI; 2001-368570/39.
XX Gene library containing sequences with specific 3'-ends and no polyA
PT tail, encoding proteins involved in a wide range of cellular processes
PT
XX
XX Claim 15; Page 38; 251pp; German.
XX
XX This invention describes a novel gene library (A) comprises a gene
CC sequence (or its part) encoding a protein involved in amino acid
CC synthesis, cellular/energy metabolism, metabolism of
CC fatty acids/phospholipids, synthesis or breakdown of
CC purines/pyrimidines/nucleosides/nucleotides, DNA
CC replication/transcription/translation, or is a transport/binding protein.
CC (A) are produced that correspond to the 3'-end of mRNA but without the
CC polyA tail. They can be prepared more efficiently and with less effort
CC than conventional libraries. AAK53436-AAK54275 represent fragments of the
XX gene library described in the method of the invention.
XX
SQ Sequence 366 BP; 99 A; 92 C; 52 G; 123 T; 0 other;

Query Match 31.7%; Score 201; DB 22; Length 366;
Best Local Similarity 78.2%; Pred. No. 3.8e-50;
Matches 279; Conservative 0; Mismatches 75; Indels 3; Gaps 3;

Qy 272 atttgacaaagtctcagatgacagaaagtcgaagcaccataacaccttagatgggg 331
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 332 tgcctggtacatgtgcagaatggatggaaatacaaccaccataaagagaaacgaga 391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 306 GGCCTGTGCGAGGTCAGAGTGGATGGAAAGTCGACCACAAATAAGAGAAACGAGA 247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 392 ggatataaactggtggtgaatgcgtcatgaaaggcgtcactccacgagagattatga 451
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 246 TGGTGACAAAGTGTGTGGTGAATGTGTATGAAAGCGTGACTCCACAAGAGTTTATGA 187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 452 gagacataaagcgaaggaactgaacctggactgaagtcgattgaactctacaacatt 511
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 186 AAGGCATGAGCCCAAGGAAG-AGCCCTGGATGGAAATTTGCATCAAACTACCAATAGT 128
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 512 ctgtgggatatattgtcaaaaagata-ttgtgttttccctgatttagcaagcaagtaa 570
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 127 CAGTCGGATTTATTCTTTTAAAGATATGATTTTCCACTAATAAGCAAGCAATTAA 68
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 571 ttttcccaagctg-atttattcaatattggtacgttgggttaataactttttt 626
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 67 TTTTCTGAAAGATGCATTTTATTGGATATGTTATGTTGATTAATAAAACCTTTT 11
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
AAC03853
ID AAC03853 standard; cDNA; 493 BP.
XX
XX AAC03853;
AC
XX
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 3851.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
PN
XX
XX 06-SEP-2000.
PD
```

```
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX P-PSDB; AAG03847.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 3851; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 493 BP; 163 A; 98 C; 120 G; 110 T; 2 other;

Query Match 30.6%; Score 194; DB 21; Length 493;
Best Local Similarity 71.5%; Pred. No. 5.5e-48;
Matches 266; Conservative 2; Mismatches 102; Indels 2; Gaps 1;

Qy 59 caaatgtgtgatgctttttagttacctggaaactgtctccagtgaaacttttgatga 118
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 116 cagatgagcaaaattccctgggcacctggaaactgtctctagtgaactttgacga 175
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 119 ttatatgaaagatagtagtgggttttggccaccaggaaagtggctggcattggcacaacc 178
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 176 ttacatgaaagctctgggtgtgggttagccaccagaaactgggaaatttggcacaacc 235
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 179 taacatgatcatcagtggaatggggatgtgacaccattaaactgaagtacacattaa 238
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 236 cactgtgatcatcagcaagaaaggagattataactatacgaactgaagtacacattaa 295
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 239 aaatactgagattcttcatactggccaggaaatttgacgaagtcactgcagatgacag 298
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 296 aaatacagaatctcttcaagtagccaggaaatttgaagaaccacagcgtgacaatag 355
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 299 gaaagtcaagagcaccataacacttagatgggggtctctctgtgtacatgtgcagaatggga 358
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 356 aaagaccaagagcatcgtaaacctgcagagagatcactgaatcgaagtgcagagatggga 415
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 359 tggaaaaatacaccaccataaagagaaacgagagatgataaactggtggtgaatgcgt 418
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 419 catgaaagcgt 430
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 474 aatgaagcgt 485
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
AAC32482
ID AAC32482 standard; DNA; 8144 BP.
XX
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 05:58:17 ; Search time 2258.96 Seconds  
(without alignments)  
5873.244 Million cell updates/sec

Title: US-09-503-596-2  
Perfect score: 634  
Sequence: 1 ggaattccaggagggtgcag.....ataacttttttagatttag 634

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_on.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
--------	-----	-------	-------------	--------	-------	-------------

1	634	100.0	634	6	AX333018
2	634	100.0	634	9	HUMALBP
3	606.6	95.7	666	9	BC003672
4	483.6	76.3	619	10	AF327855
5	438	69.1	625	4	BTATFABP
6	417.8	65.9	573	4	SSC416020
7	386.2	60.9	614	10	MUSLBP
8	341.8	53.9	600	10	AF144756
9	341.6	53.9	542	10	MUS13KMP
10	313.4	49.4	395	6	A98023
11	298.2	47.0	399	10	RNU75581
12	295.8	46.7	339	4	AF136241
13	249.6	39.4	296	4	AF102872
14	240	37.9	472	5	AF432506
15	232	36.6	1836	4	RABPLP2
16	215.8	34.0	2150	9	HSP2RNA
17	211.4	33.3	411	6	A92933
18	211	33.3	165948	9	AC018616
19	211	33.3	174404	2	AC013542
20	203	32.0	374	6	AX046512
21	174.2	27.5	561	10	RNU07870
22	174.2	27.5	581	10	RNU09022
23	171.2	27.0	666	10	RATFABPH
24	170.8	26.9	399	10	MMU96149
25	169.6	26.8	652	10	RATFABPHA
26	168.4	26.6	638	4	SSC416019
27	167.6	26.4	658	10	MMHFABP
28	166.2	26.2	701	10	BC002082
29	165.4	26.1	688	10	AF327854
30	165.4	26.1	879	9	BC007021
31	164.6	26.0	1778	10	MMU02885
32	164	25.9	669	10	MMU02883
33	164	25.9	705	5	AF237712
34	163.8	25.8	551	9	HSFAB
35	162.6	25.6	671	4	BTU55188
36	162.6	25.6	677	4	AF059507
37	162.4	25.6	671	4	BTMDGI
38	161.2	25.4	640	9	HSMDGI
39	160.8	25.4	684	4	BTFABPH
40	159.8	25.2	654	5	AF448057
41	159.4	25.1	8144	4	SSY16039
42	159.4	25.1	8144	6	A98022
43	158.8	25.0	695	10	S83247
44	157.8	24.9	313	6	AX046511
45	157.2	24.8	664	10	RNU13253

ALIGNMENTS

RESULT	1
AX333018	
LOCUS	AX333018 634 bp DNA linear PAT 09-JAN-2002
DEFINITION	Sequence 3527 from Patent WO0194629.
ACCESSION	AX333018
VERSION	AX333018.1 GI:18123652
KEYWORDS	human.
SOURCE	
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
	Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE	Cancer gene determination and therapeutic screening using signature
JOURNAL	gene sets
Patent:	WO 0194629-A 3527 13-DEC-2001;
FEATURES	Avalon Pharmaceuticals (US)
source	Location/Qualifiers
	1..634
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	200 a 108 c 151 g 175 t

ORIGIN

Query Match 100.0%; Score 634; DB 6; Length 634;  
Best Local Similarity 100.0%; Pred. No. 4e-153;  
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggaattccaggagggtgcagcttccttcacacttgaaagaataaactcagaactcaca 60  
DB 1 GGAATTCCAGGAGGGTGCAGCTTCCTTCACCTGCAAGAATAAATCCTAGAAAACCTCACA 60

QY 61 aaatgtgtgatgcttttttagtagtaactgaaactgtctccagtgaaaaactttgatgatt 120  
DB 61 AAATGTGTGATGCTTTTGTAGTAGTACCTGGAACACTTGTCTCCAGTGAACACTTTGTGATGATT 120

QY 121 atataaagaagaatagagtggtgctttgcccaggaaagtgcctggcagcagcacaacta 180  
DB 121 ATATGAAGAAGTAGAGTGGGCTTTGCCAGGAAAGTGGCTGGCATGGCCAAACCTTA 180

QY 181 acatgatcatcagtggaatgggagtgatcaccattaaatctgaaagtacactttaaaa 240  
DB 181 ACATGATCATCAGTGTGAATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTTAAA 240

QY 241 atactgagattcccttcatactggtggcaggaatttgacagatcagatgacagga 300  
DB 241 ATACTGAGATTTCCTTCATCTACTGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGA 300

QY 301 aagtcagaagcaccataaacttagatgggggtgctctgtacatgtgcagaatgggagtg 360  
DB 301 AAGTCAAGAGCACCATAAACCCTTAGATGGGGGTGCTCTGTGATGTGCAAGAAATGGGATG 360

QY 361 gaaatcaaccaccataaagaagaacagagagatgataaactggtgggaatgcgtca 420  
DB 361 GAAATCAACCACCATAAAGAGAAAACGAGAGGATGATAAATGGTGGTGAATGCGTCA 420

QY 421 tgaaggcgtcactccagagagtttatgagagagcataaagcgaaggaactgacctg 480  
DB 421 TGAAGGCGTCACTTCCACGAGAGTATGAGAGAGCATAAAGCAAGGACGTTGACCTG 480

QY 481 gactgaagtgcattgaactcacaacttctggtgatatattgttcaaaaagatat 540  
DB 481 GACTGAAGTTCGATGAACCTACACATCTGTGGGATATATGTTCAAAAAGATATT 540

QY 541 gtgttttccctgattagcaagcaagtaattttcccaagctgattttatcattatg 600  
DB 541 GTGTGTTTCCCTGATTAGCAAGCAAGTAATTTTCTCCCAAGCTGATTTTATTCAATATG 600

QY 601 gttacgttggttaataaacttttttttagatttag 634  
DB 601 GTTACGTTGGTTAAATAACTTTTTTTTAGATTTAG 634

RESULT 2

LOCUS HUMALBP 634 bp mRNA linear PRI 31-OCT-1994  
DEFINITION Human adipocyte lipid-binding protein, complete cds.  
ACCESSION J02874  
VERSION J02874.1 GI:178346  
KEYWORDS adipocyte lipid-binding protein.  
SOURCE Human adipose, cDNA to mRNA, clone lambda-BALBP.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 634)  
Baxa,C.A., Sha,R.S., Buelt,M.K., Smith,A.J., Matarese,V.,  
Chinander,L.L., Boundy,K.L. and Bernlohr,D.A.  
Human adipocyte lipid-binding protein: purification of the protein  
and cloning of its complementary DNA  
Biochemistry 28 (22), 8683-8690 (1989)  
90105397  
Draft entry and computer-readable copy of sequence [1] kindly  
submitted R.Sha, 07-AUG-89.  
Location/Qualifiers

source 1..634  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="Unassigned"  
1..634  
/gene="FABP4"  
63..461  
/gene="FABP4"  
/note="adipocyte lipid-binding protein"  
/codon\_start=1  
/protein\_id="AA51689.1"  
/db\_xref="GI:178347"  
/db\_xref="GDB:G00-128-030"  
/translation="MCDAFVGTWKLVSSSENFDDYMKVEVGVGFATRKVAGMAKPNMIIS  
VNGDVITIKSESTFKNTETISILGQEFDEVTAADKRKVKSTITLDGGLVHVQKWDGKS  
TTIRKREDDKLVVECVMKGVSTRYERA"  
polyA\_signal 613..618  
/gene="FABP4"  
BASE COUNT 200 a 108 c 151 g 175 t  
ORIGIN 2 bp upstream of EcoRI site.

Query Match 100.0%; Score 634; DB 9; Length 634;  
Best Local Similarity 100.0%; Pred. No. 4e-153;  
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggaattccaggagggtgcagcttccttcacacttgaaagaataaactcagaactcaca 60  
DB 1 GGAATTCCAGGAGGGTGCAGCTTCCTTCACCTTGAAGAATAAATCCTAGAAAACCTCACA 60

QY 61 aaatgtgtgatgcttttttagtagtaactgaaactgtctccagtgaaaaactttgatgatt 120  
DB 61 AAATGTGTGATGCTTTTGTAGTAGTACCTGGAACACTTGTCTCCAGTGAACACTTTGTGATGATT 120

QY 121 atataaagaagaatagagtggtgctttgcccaggaaagtgcctggcagcagcacaacta 180  
DB 121 ATATGAAGAAGTAGAGTGGGCTTTGCCAGGAAAGTGGCTGGCATGGCCAAACCTTA 180

QY 181 acatgatcatcagtggaatgggagtgatcaccattaaatctgaaagtacactttaaaa 240  
DB 181 ACATGATCATCAGTGTGAATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTTAAA 240

QY 241 atactgagattcccttcatactggtggcaggaatttgacagatcagatgacagga 300  
DB 241 ATACTGAGATTTCCTTCATCTACTGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGA 300

QY 301 aagtcagaagcaccataaacttagatgggggtgctctgtacatgtgcagaatgggagtg 360  
DB 301 AAGTCAAGAGCACCATAAACCCTTAGATGGGGGTGCTCTGTGATGTGCAAGAAATGGGATG 360

QY 361 gaaatcaaccaccataaagaagaacagagagatgataaactggtgggaatgcgtca 420  
DB 361 GAAATCAACCACCATAAAGAGAAAACGAGAGGATGATAAATGGTGGTGAATGCGTCA 420

QY 421 tgaaggcgtcactccagagagtttatgagagagcataaagcgaaggaactgacctg 480  
DB 421 TGAAGGCGTCACTTCCACGAGAGTATGAGAGAGCATAAAGCAAGGACGTTGACCTG 480

QY 481 gactgaagtgcattgaactcacaacttctggtgatatattgttcaaaaagatat 540  
DB 481 GACTGAAGTTCGATGAACCTACACATCTGTGGGATATATGTTCAAAAAGATATT 540

QY 541 gtgttttccctgattagcaagcaagtaattttcccaagctgattttatcattatg 600  
DB 541 GTGTGTTTCCCTGATTAGCAAGCAAGTAATTTTCTCCCAAGCTGATTTTATTCAATATG 600

QY 601 gttacgttggttaataaacttttttttagatttag 634  
DB 601 GTTACGTTGGTTAAATAACTTTTTTTTAGATTTAG 634

RESULT 3  
BC003672



LOCUS BC003672 666 bp mRNA linear PRI 12-JUL-2001  
 DEFINITION Homo sapiens, fatty acid binding protein 4, adipocyte, clone  
 MGC:12293 IMAGE:3683235, mRNA, complete cds.  
 ACCESSION BC003672  
 VERSION BC003672.1 GI:13277535  
 KEYWORDS MGC.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 666)  
 Strausberg,R.  
 Direct Submission  
 Submitted (28-FEB-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 16 Row: m Column: 7  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4557578.

FEATURES  
 source  
 1..666  
 /organism="Homo sapiens"  
 /db\_xref="LocusID:2167"  
 /db\_xref="taxon:9606"  
 /clone="MGC:12293 IMAGE:3683235"  
 /tissue\_type="Bladder, carcinoma"  
 /clone\_lib="NIH\_MGC\_53"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"  
 67..465

## CDS

/product="fatty acid binding protein 4, adipocyte"  
 /protein\_id="AAH03672.1"  
 /db\_xref="GI:13277536"  
 /translation="MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIIS  
 VNGDVTIKSESTFKNTEISFILGOEFDVADRRKVKSTITLDGGVLVHVORWGDGKS  
 TTIKRREDDKLVECVMKGYSTRVERA"  
 BASE COUNT 232 a 116 c 144 g 174 t  
 ORIGIN

Query Match 95.7%; Score 606.6; DB 9; Length 666;  
 Best Local Similarity 98.6%; Pred. No. 4.7e-146;  
 Matches 612; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 6 tcacagagggtgcagcttccttcaccttgagataatcctgaaataatcctgaaataatcctgaaataatc 65  
 Db 10 TCCTGAAACTGCAGCTTCCTCACCCTGGAAGTAATATCTAGAAAACCTCACAAGATG 69  
 Qy 66 tctgatgctttgttaggtacacgttgaaactgtctcagtgaaactgttgatgattatg 125  
 Db 70 TGTGATGCTTTGTAGGTACCTGGAACCTGTCTCCAGTGAAACTTTGATGATTATG 129  
 Qy 126 aaagaagttaggtggcttggccaccaggaagtgagctggcgtgacacacctaactg 185  
 Db 130 AAAGAAGTAGGAGTGGGGTTTGGCCACCAGGAAAGTGGCTGGCATGGCCAAACCTAACATG 189

Qy 186 atcatcagtgtaattggggatgtgatcaccattaaatcctgaaagtaccttttaaaataact 245  
 Db 190 ATCATCAGTGTGAATGGGATGTGATCACCATTAAATCTGAAAGTACCTTTAAAAATACT 249  
 Qy 246 gagatttccttcatactggtggccaggaaatttgacgaagtccactgcagatgacacaggaaagtc 305  
 Db 250 GAGATTTCCTTCATCTGCGCCAGGAATTTGACGAAGTCTCAGATGCAGATGCACAGGAAGTC 309  
 Qy 306 aagagcaccaataaccttagatgggggtgtcctcgtgtacatgtgcagaaatgggatgaaata 365  
 Db 310 AAGAGCACCATTAACCTTAGATGGGGTGTCTCGTACATGTGCAGAAATGGGATGAAAA 369  
 Qy 366 tcacacacataaaacagaaacagagagatgataaactgtgtgtggaatgcgtcatgaaa 425  
 Db 370 TCACACCATTAAGAGAAAAACGAGAGATGATAAACTGGTGTGGAAATGCGGTCAATGAA 429  
 Qy 426 ggctcacttcacagagagatttatgagagagcataagccaagggagcgttgacctgactg 485  
 Db 430 GCGTCACTTCACGAGAGTTTATGAGAGAGCATAGCCCAAGGGAGCTTGACCTGGACTG 489  
 Qy 486 aagttcgatgaaactctacaactctctgtgggatatattgttcaaaaagatattgtgt 545  
 Db 490 AAGTTGCGATTGAACCTCTACAACATTTCTGTGGATATATTGTTCAAAAAGATATTGTTGT 549  
 Qy 546 ttctcctgatttagcaagaagaataattttctccaaagctgattttatcaatatgtttac 605  
 Db 550 TTTCCATGATTAGCAAGCAACTAATTTTCTCCCAAGCTGATTTTATTAATCAATGTTTAC 609  
 Qy 606 gtgggttaataaactttttt 626  
 Db 610 GTTGGTAAATAACTTTT 630

RESULT 4  
 LOCUS AF327855 619 bp mRNA linear ROD 14-JAN-2002  
 DEFINITION Sperophilus tridecemlineatus adipose-type fatty acid-binding  
 protein (FABP4) mRNA, complete cds.  
 ACCESSION AF327855  
 VERSION AF327855.1 GI:12802819  
 KEYWORDS thirteen-lined ground squirrel.  
 SOURCE Sperophilus tridecemlineatus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
 Sperophilus.  
 REFERENCE 1 (bases 1 to 619)  
 AUTHORS Hittel,D. and Storey,K.B.  
 TITLE Differential expression of adipose- and heart-type fatty acid  
 binding proteins(1) in hibernating ground squirrels  
 JOURNAL Biochim. Biophys. Acta 1522 (3), 238-243 (2001)  
 MEDLINE 21638064  
 PUBMED 11779641  
 REFERENCE 2 (bases 1 to 619)  
 AUTHORS Hittel,D.S. and Storey,K.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-DEC-2000) Biology, Carleton University, 1125 Colonel  
 By Drive, Ottawa, ON K1S 5B6, Canada  
 FEATURES  
 Location/Qualifiers  
 1..619  
 /organism="Sperophilus tridecemlineatus"  
 /db\_xref="taxon:43179"  
 1..619  
 /gene="FABP4"  
 47..445  
 /gene="FABP4"  
 /note="differentially up-regulated in hibernating versus  
 euthermic squirrels; cytoplasmic"  
 /codon\_start=1  
 /product="adipose-type fatty acid-binding protein"  
 /protein\_id="AAK08084.1"  
 /db\_xref="GI:12802820"  
 /translation="MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIIS



```
RESULT 6
SSC416020          573 bp      mRNA      linear      MAM 21-OCT-2001
LOCUS              Sus scrofa partial mRNA for adipocyte fatty acid-binding protein
DEFINITION          (FABP4 gene).
ACCESSION           AJ416020
VERSION             AJ416020.1 GI:16304809
KEYWORDS             adipocyte fatty acid-binding protein; FABP4 gene.
SOURCE              pig.
ORGANISM             Sus scrofa
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE            1 (bases 1 to 573)
AUTHORS              Gerbens,F.
TITLE                Genetic control of intramuscular fat accretion in pigs. The role of
                    heart and adipocyte fatty acid-binding proteins
JOURNAL              Thesis (2000) Department of Biochemistry MW, University of
                    Nijmegen, Nijmegen, Netherlands
REFERENCE            2 (bases 1 to 573)
AUTHORS              Gerbens,F., Jansen,A., van Erp,A.J., Harders,F., Meuwissen,T.H.,
                    Rettenberger,G., Veerkamp,J.H. and de Pas,M.F.
TITLE                The adipocyte fatty acid-binding protein locus: characterization
                    and association with intramuscular fat content in pigs
JOURNAL              Mamm. Genome 9 (12), 1022-1026 (1998)
MEDLINE              99099248
AUTHORS              Gerbens,F.
TITLE                Direct Submission
JOURNAL              Submitted (04-OCT-2001) Gerbens F., Animal Sciences, ID-Lelystad,
                    P.O.Box 65, NL-8200AB, NETHERLANDS
FEATURES             location/Qualifiers
                    1..573
                    /organism="Sus scrofa"
                    /db_xref="taxon:9823"
                    1..399
                    /gene="FABP4"
                    <1..399
                    /gene="FABP4"
                    /function="intracellular fatty acid-binding"
                    /citation=[2]
                    /codon_start=1
                    /product="adipocyte fatty acid-binding protein"
                    /protein_id="CAC95166.1"
                    /db_xref="GI:16304810"
                    /translation="MCDAFVGTWKLVSSEDDYMKVGVGFATRKVAGMAKPNLIIF
                    VNGDMITIRSESTFKNTEIAFKLQGEDEVTDADRKKVSTITLDGALVQVKWDGKT
                    TTINKIYDVKLVVEICIMKGVATRIVERA"
                    537..559
                    /gene="FABP3"
                    /gene="FABP3"
                    /gene="FABP3"
                    /evidence-experimental
BASE COUNT          194 a 95 c 131 g 152 t 1 others
ORIGIN
Query Match          65.9%; Score 417.8; DB 4; Length 573;
Best Local Similarity 85.8%; Pred. No. 2.4e-97;
Matches 490; Conservative 1; Mismatches 68; Indels 12; Gaps 2;
QY 63 atgtgtgatcgtttgttagtaccctggaaactgtctccagtgaaactttgatgattat 122
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATGTGTGATGCATTCTAGTAGTACCTGGAACTTGCTCCAGTGAAACTTTGATGATTAC 60
QY 123 atgaagaagtggagtggtggttgcaccaggaagtggctggcagccaaactaac 182
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ATGAAGAAGTGGGTGGCTTTGCCACAGGAAGTGGCTGGCAGCCAAACCAAC 120
QY 183 atgatcatcagtgtaagtgggatgtgatcaccattaaatctgaaagtaccctttaa 242
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 121 CTGATCATCACTGTGAATGGGATATGATCACCATTAGATCAGAAAGTACCTTTAAAAAT 180
QY 243 actgagatttcctcattctggccagggaatttgacgaagtcactgcagatgacaggaaa 302
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ACTGAGATTGCCTTCAAAATTTGGCCAGGAATTTGATGAAGTCACTGCAGATGACAGGAAA 240
QY 303 gtcaagagaccaccataacaccttagatgggggtgctcctgggtacatggtgcagaaaatgggatgga 362
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GTCAAGAGCACCATAACCTTAGTAGGAGGCGCCTGGTACAGTGCAGAAAGTGGGATGGA 300
QY 363 aaatcaaccaccataaagaaacgagagatgataaactggtggtggaatgcgtcgt 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AAGACACACCACCCTAAACAGAAAAATAGTAGATGATAGTCTGGTGGTGGAAATGATATCATG 360
QY 423 aaagcgtcactccacagagagtttatgagagagcataagcgaagggacgttgacctgga 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AAAGGTGTACGGCTTACCAGAAATTTATGAGAGAGCATAAGCCA-----AATTGTA 410
QY 483 ctgaagttcgatgaactctcaaacattctgtgggatatattgttcaaaaagatatgtc 542
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 411 CARAAGTTTGGATCGAACTCTACAACACTCTGTGTGGATGATATTGCCGACATATGTTGT 470
QY 543 ttttccctgatttagcaagcaagttaatttctcccaagctgattttcaaatatgt 602
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 471 TATTTCCACTAATTAGGAGCAACTAATTTTCCCCACACATGATTTTCTTCAATATGCT 530
QY 603 tacgttggttaataaacttttttagattta 633
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 531 T--GTTAGTTAAATAAAACTTTTATGATTTA 559
RESULT 7
MUSLBP             MUSLBP             614 bp      mRNA      linear      ROD 12-JUN-1993
LOCUS              Mouse 3T3-L1 lipid binding protein mRNA, complete cds.
DEFINITION          K02109
ACCESSION           K02109
VERSION             K02109.1 GI:198716
KEYWORDS             lipid binding protein.
SOURCE              Mouse 3T3-L1 adipocytes, cDNA to mRNA, clone pAL422.
ORGANISM             Mus musculus
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE            1 (bases 1 to 614)
AUTHORS              Bernalhr,D.A., Angus,C.W., Lane,M.D., Bolanowski,M.A. and
                    Kelly,T.J.
TITLE                Expression of specific mRNAs during adipose differentiation:
                    Identification of mRNA encoding a homologue of myelin P2 protein
JOURNAL              Proc. Natl. Acad. Sci. U.S.A. 81, 5468-5472 (1984)
MEDLINE              84298159
COMMENT              The sequence of pAL422 (422 protein) is 69% homologous to rabbit
                    myelin P2 protein and 64% homologous to bovine myelin P2 protein,
                    which suggests that 422 protein is an analog of myelin P2.
                    Additional homologues with fatty acid binding proteins from rat
                    liver and rat intestine suggest that 422 protein may be a lipid
                    transport protein in adipocytes, and myelin P2 proteins may
                    function similarly in Schwann cells.
FEATURES             Location/Qualifiers
                    1..614
                    /organism="Mus musculus"
                    /db_xref="taxon:10090"
                    <1..614
                    /gene="L1"
                    /product="LBP mRNA"
                    1..614
                    /gene="L1"
                    44..442
                    /gene="L1"
                    /codon_start=1
                    /product="lipid-binding protein"
                    /protein_id="AAA39416.1"
                    /db_xref="GI:293695"
                    /translation="MCDAFVGTWKLVSSEDDYMKVGVGFATRKVAGMAKPNMIIS
                    VNGDLVTIRSESTFKNTEISFKLGVFEITADRRKVKSIITLDGALVQVKWDGKS
```



ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 542)
AUTHORS	Cook, K.S., Hunt, C.R. and Spiegelman, B.M.
TITLE	Developmentally regulated mRNAs in 3T3-adipocytes: Analysis of transcriptional control
JOURNAL	J. Cell Biol. 100, 514-520 (1985)
MEDLINE	85105214
FEATURES	Location/Qualifiers
source	1..542
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/cell_line="3T3"
	/cell_type="adipocyte"
mRNA	<1..542
	/product="13k protein"
CDS	<1..369
	/codon_start=1
	/product="13K protein"
	/protein_id="AAA37112.1"
	/db_xref="GI:191493"
translation	"LVSSNFDDYMEVGFATRVKAGMAKPNMIISVNGDLVTTRS ESTFKNTEISFLGVDFEFTADDRKVKSIITLDGALYOVQKWGDGSTTKRRDGD KLVECEVMKVSTSTRYERA"
BASE COUNT	173 a    90 c    140 g    139 t
ORIGIN	
	Query Match                  53.9%; Score 341.6; DB 10; Length 542;
	Best Local Similarity        81.0%; Pred. No. 1.1e-77;
	Matches 434; Conservative    0; Mismatches 99; Indels 3; Gaps 3;
QY	93 ctgtctccagtgaaacttggattatataaaaagaagttagagtgggcctttgcccacc 152
Db	1 CTGTCTCCAGTGAACACTTCGATGATTACATGAAGAAGTGGAGTGCGCTTTGCCACA 60
QY	153 agaaaggcgctggcatggccaaacctacaatgatcatcagtgtaatggggatgtagtc 212
Db	61 AGAAAGTGGCAGCATGCGCACAGCCCAACATGATCATCAGCGTAATAATGGGGATTTGGTC 120
QY	213 accattaaatctgaagtaacctttaaaaaactagtagattccttcatactagggccaggaa 272
Db	121 ACCATCCGGTCAGAGAGTACTTTTAAAAAACACCAGAGATTTCCRTCAAACCTGGCGTGGAA 180
QY	273 tttagcaagaactgcactgcagatgacaggaagaagtccaagagcacccataaaccttagatgggggt 332
Db	181 TTCGATGAATATCACCGCAGACGACAGGAAGGTGAAGAGCATCATACCCTAGATGGCGGG 240
QY	333 gtctcgttacatgtgcagaatatggatggaaatcaaccaccataaagagaaaaacagagag 392
Db	241 GCCTGTGTGCAGGTGCACAGAGTGGGATGGAAAGTCGACCACCATTAAGAGAGAAAACGAGAT 300
QY	393 gatgataaacgtgtgtgggaatgogtcatgaaagggcgtcacttcacagagagtttatgag 452
Db	301 GGTCACAAGCTGGTGTGGAATGTGTTATGAAGGGCTGACTTCCACAAGAGTTTATGAA 360
QY	453 agagcataagccaaggacgttgacctggactgaagtcgaattgcattgaactacaaatttc 512
Db	361 AGGCATATGAGCCAAAGGAAG-AGGCCCTGGATGGAAATTTGCATCAAAACATACAAATAGTC 419
QY	513 tgtgggataattgttcaaaaaagata-ttgtftgttttccctgatttagcaagcaaglaat 571
Db	420 AGTCGGATTATTGTTTTTTTTTAAAGATATGATTTTCCACTAATATAGCAAGCAATTAAT 479
QY	572 ttctcccaagctg-atatttattoaatatggttcacgttggttaaataactttttt 626
Db	480 CTTTCTCTGAAGATGCATTTATTGGATATGGTATGTTATGTTATTAATAAAACCTTTT 535

RESULT	10				
A98023					
LOCUS	A98023	395 bp	DNA	linear	PAT 26-JAN-2000

Sequence 2 from Patent WO9914365.									
DEFINITION	A98023								
ACCESSION	A98023.1								
VERSION	GI:6781261								
KEYWORDS	unidentified.								
SOURCE	unidentified.								
ORGANISM	unclassified.								
REFERENCE	1 (bases 1 to 395)								
AUTHORS	Gerbens,F.								
TITLE	THE PORCINE ADIPOCYTE FATTY ACID-BINDING PROTEIN ENCODING GENE AND METHODS TO LOCALISE, IDENTIFY OR MARK GENES OR ALLELES OR QUANTITATIVE TRAIT LOCI OF FARM ANIMALS								
JOURNAL	Patent: WO 9914365-A 2 25-MAR-1999;								
FEATURES	STAMBOEK ZUID B V (NL); DALLAND B V (NL)								
source	Location/Qualifiers								
	1..395								
	/organism="unidentified"								
	/db_xref="taxon:32644"								
BASE COUNT	132 a	67 c	105 g	91 t					
ORIGIN									
	Query Match 49.4%; Score 313.4; DB 6; Length 395;								
	Best Local Similarity 90.0%; Pred No. 2.1e-70;								
	Matches 359; Conservative 0; Mismatches 36; Indels 4; Gaps 2;								
Qy	63	atgtgtgatgctttttaggtaccttgaaacctgtctccagtgaaaccttgatgattat	122						
Db	1	ATGTGTGATGCATTTGTAGGTACCT-GAAACTTGTCTCCAGTGAAAACCTTTGATGATTAC	59						
Qy	123	atgaaaagatagagatggcctttgcacacaggaagaagtgcctggcatggccaaacctaac	182						
Db	60	ATGAAAAGATGGGAGTGGGCTTTGCCACCAGGAAGTGGCTGGCATGGCCAAACCCAAC	119						
Qy	183	atgatcatcagtgtgaatgggagtgtgaccattaaactgaaagtacatttcaaaaat	242						
Db	120	CTGATCATCACTGTGATGGGATATGATCACCATTAGATCAGAAAGTACCTTTAAAAAT	179						
Qy	243	actgagatttctcttactcaggccaggaatttgacaaagtcactgcagatgacagaaa	302						
Db	180	ACTGAGATTGCTTCAAATTTGGCCAGGAATTTGATGAAGTCACTGCAGATGACAGAAA	239						
Qy	303	gtcaagagaccataaaccttagatgggggtgtcctgtacatgtgcagaaatgggatgga	362						
Db	240	GTCAAGAGCACCAATCACTTAGATGGAGGCCCTGTTACAGGTGCAGAGTGGGATGGA	299						
Qy	363	aaatcaaccaccataaagagaaaacgagagatgataaaactggtgggaatgcgtcatg	422						
Db	300	AAGACACCACCAATACACAGAAAAATA--GTAGATAAGTTGGTGGTGAATGTATCATG	356						
Qy	423	aaaggctcacttccacgagagatttatgagagacataa	461						
Db	357	AAAGGTGTACGGGTACCAAGATTTATGAGAGAGCATAA	395						

RESULT	11
LOCUS	RN075581
DEFINITION	399 bp mRNA linear ROD 23-MAR-2001 Rattus norvegicus adipocyte fatty acid binding protein (A-FABP)
ACCESSION	U75581
VERSION	U75581.1
KEYWORDS	GI:1658524
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 399)
AUTHORS	Prinsen,C.F. and Veerkamp,J.H.
TITLE	Transfection of L6 myoblasts with adipocyte fatty acid-binding protein cDNA does not affect fatty acid uptake but disturbs lipid metabolism and fusion

JOURNAL Biochem. J. 329 (Pt 2), 265-273 (1998)  
MEDLINE 98087494  
PUBMED 9425108  
REFERENCE 2 (bases 1 to 399)  
AUTHORS Prinsen,C. and Veerkamp,J.  
TITLE Direct Submission  
JOURNAL Submitted (22-OCT-1996) Dept. Biochemistry, University of Nijmegen,  
Adelbertusplaats 1, Nijmegen 6500 HB, The Netherlands  
FEATURES Location/Qualifiers  
source 1..399  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
1..399  
/gene="A-FABP"  
1..399  
/gene="A-FABP"  
/codon\_start=1  
/product="adipocyte fatty acid binding protein"  
/protein\_id="AAB18344.1"  
/db\_xref="GI:1658525"  
/translation="MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNLIIS  
VEGLDVIRSESTFKNTEISFKLGVEFDEITPDDRKVKSIITLDGGVLVHVQKWDGKS  
TTIKRRDGDKLVECYMKGVTSTVVERA"  
BASE COUNT 118 a 73 c 119 g 89 t  
ORIGIN  
Query Match 47.0%; Score 298.2; DB 10; Length 399;  
Best Local Similarity 84.2%; Pred. No. 1.8e-66;  
Matches 336; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 63 atgtgtgatgtctttttaggtacctggaaacttctccagtgaaacttttgatgattat 122  
Db 1 ATGTGTGATGCTTTGTGGGACCTGGAAACTCGCTCCAGTGAGAACTTCGATGATTAC 60  
QY 123 atgaagaagttaggtgggttttccaccaggaagaagtggctggcatggccaaacctaac 182  
Db 61 ATGAAGAAGTGGGAGTGGCTTCCGCCACCAGGAAGTGGCCGGTATGCCAAGCCCAAC 120  
QY 183 atgatcatcagtgtaagggtgatgtatcaccattaaactctgaaagtacaccttttataat 242  
Db 121 TTGATCATCAGGTAGAGGGGACTTGGTCGTCATCCGGTCAGAGAGTACTTTTAAAC 180  
QY 243 actgagattcttcatactactggccaggaatttgacgaagtcactgcagatgacaggaaa 302  
Db 181 ACCGAGATTTCTTCAAACTGGGTGTGGAATTCGATGAAATCACCCAGATGACAGGAAA 240  
QY 303 gtcaagagcacataacctatagatgggtgtcctcgtgtacatgtgcagaaatgggatgga 362  
b 241 GTGAAGAGCATCATAACTCGATGGTGGGTGCTCTGGTACATGTGCAAGTGGGATGGA 300  
QY 363 aaatcaacacacataaagagaaacagagagatgataaactgggtggaatgcgtcatg 422  
Db 301 AAGTCGACCACCATAAAGAGAGACGAGATGGTGACAAGCTGGTGGTGAATGTGTCATG 360  
QY 423 aaagcgcacttcacacagagagtttatgagagagacataa 461  
Db 361 AAAGCGTGACTTCTACAAGAGTATTATGAAAGAGCATAA 399  
RESULT 12  
AF136241  
LOCUS AF136241 339 bp mRNA linear MAM 25-MAY-1999  
DEFINITION Oryctolagus cuniculus adipocyte lipid-binding protein (ap2) mRNA,  
partial cds.  
ACCESSION AF136241  
VERSION AF136241.1 GI:4887136  
KEYWORDS rabbit.  
SOURCE Oryctolagus cuniculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
REFERENCE 1 (bases 1 to 339)  
AUTHORS Guan,Y., Zhang,Y., Davis,L., Davis,L. and Breyer,M.D.  
TITLE Expression of ap2 gene in transitional epithelium of rabbit bladder  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 339)  
AUTHORS Guan,Y., Zhang,Y., Davis,L. and Breyer,M.D.  
TITLE Direct Submission  
JOURNAL Submitted (19-MAR-1999) Division of Nephrology, Vanderbilt  
University Medical Center, 21st street, Nashville, TN 37232-2372,  
USA  
FEATURES Location/Qualifiers  
source 1..339  
/organism="Oryctolagus cuniculus"  
/strain="New Zealand White"  
/db\_xref="taxon:9986"  
/tissue\_type="bladder"  
<1..>339  
/gene="ap2"  
/gene="ap2"  
/note="AFABP"  
/codon\_start=1  
/product="adipocyte lipid-binding protein"  
/protein\_id="AAD32209.1"  
/db\_xref="GI:4887137"  
/translation="TWKLVSSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGDVIT  
IKSESTFKNTEISFKLGVEFDEITADDRKKSIIITLDGGALVQVQKWDGKSTTIKRR  
EGDKLVVECYM"  
BASE COUNT 108 a 63 c 97 g 71 t  
ORIGIN  
Query Match 46.7%; Score 295.8; DB 4; Length 339;  
Best Local Similarity 92.0%; Pred. No. 7.3e-66;  
Matches 312; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
QY 84 acctggaactgtctccagtgaaactttgatgatttatataagaagaagtagtgaggc 143  
Db 1 ACCTGGAAGCTGTCTCCAGTGAGAACTTTCAGCATATACATGAAAAGAGTGGGAGTGGGC 60  
QY 144 ttgcccaccaggaagtggctggcatggccaaaccttaacatgacatgacatgagtgatgag 203  
Db 61 TTCGCCACCAGGAAGTGGTGGCATGGCCAAACCCACATGATCATCATGATGATGAGTGG 120  
QY 204 gatgtgataccataataatctgaaagtacacctttaaataactagatttccttcatactg 263  
Db 121 GATGTGATCACCATTAAATACAGAGAGCACCTTTAAAAACACACTGACATTTCTTCAAACTG 180  
QY 264 gcccaggaatttgacgaagtcactgcagatgacagaaagtcacagaccataaccctta 323  
Db 181 GCCCAGGAATTTGATGAAGTACCAGATGACAGGAAGGCAAGAGTATCATATAACCTTA 240  
QY 324 gatgggggtgctcctgtgtacatgtgcagaaatgggagtggaatgcgtcatg 422  
Db 241 GATGGTGGTCCCTGGTCCAGGTGCAGAAAGTGGGATGGGAATAATCAACCCACATAAGAGA 300  
QY 384 aaacagagagatgataaactgggtggaatgcgtcatg 422  
Db 301 AAACGGGAGGTGATAAACTGGTGGTGAATGTGTTATG 339  
RESULT 13  
AF102872  
LOCUS AF102872 296 bp mRNA linear MAM 12-APR-2000  
DEFINITION Sus scrofa adipocyte fatty acid binding protein (ap2) mRNA, partial  
cds.  
ACCESSION AF102872  
VERSION AF102872.1 GI:3907628  
KEYWORDS pig.  
SOURCE Sus scrofa  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 296)

AUTHORS Ding S.T., McNeel, R.L. and Mersmann, H.J.  
TITLE Expression of porcine adipocyte transcripts: tissue distribution and differentiation in vitro and in vivo  
JOURNAL Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 123 (3), 307-318 (1999)  
MEDLINE 99410873  
PUBMED 10481259  
REFERENCE 2 (bases 1 to 296)  
AUTHORS Ding S.T. and Mersmann, H.J.  
TITLE Direct Submission  
JOURNAL Submitted (29-OCT-1998) Pediatrics, Baylor College of Medicine, 1100 Bates St., Houston, TX 77030, USA  
FEATURES Location/Qualifiers  
source 1..296  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
gene <1..>296  
CDS /gene="ap2"  
<1..>296  
/gene="ap2"  
/note="cytosolic protein"  
/codon\_start=1  
/product="adipocyte fatty acid binding protein"  
/protein\_id="AAC78684.1"  
/db\_xref="GI:3907629"  
/translations="GPATRKVAGMAKPNLITVNGDMITIRSESTFKNTEIAFKLGOE FDEVTADDRKVKSTIILDGALVQVQKWDGKTTINRKIVDDKLVECVMGVTA"  
BASE COUNT 98 a 54 c 79 g 65 t  
ORIGIN

Query Match 39.4%; Score 249.6; DB 4; Length 296;  
Best Local Similarity 90.2%; Pred. No. 6.1e-54;  
Matches 267; Conservative 0; Mismatches 29; Indels 0; Gaps 0;  
QY 141 ggccttgcacacgaggaagtgcctggcattgcccacacccaacccaacccatgacatcagtgatgaat 200  
Db 1 GCGTTTGCTACCCAGGAAGTGGCTGGCATGGCCAAACCCAAACCTGATCATCACGTGGAAT 60  
QY 201 ggggatgtgacacattaaatcctgaaagtacacctttaaataactgagattccttcata 260  
Db 61 GGGGATATGATCACCATTAGATCAGAAAGTACCTTTAAATAACTGAGATTGCCCTTCAAA 120  
QY 261 ctggccagggaattgacgaagtcaactgcagatcacaggaagaagcaagcaccataaac 320  
Db 121 TTGGCCAGGAATTTGATGAGTCACTCCAGATGACAGAAAGTCAAGAGCACCATAACC 180  
QY 321 ttgatgggggtgctcctgtgacatgtgcagaaatggggtggaataatcaaccaccataaag 380  
Db 181 TTAGATGGAGCGCCCTGGTACAGTGCAGAAAGTGGGATGCAAGACACACCATAAAC 240  
QY 381 agaaacagagaggatgataaactgggtggaatgcgtcatgaaaggcgtcacttc 436  
Db 241 AGAAAAATAGTAGATGATGAAGTTGGTGGTGAATGTGTGTCATGAATGTGTCACTGC 296

RESULT 14  
AF432506 472 bp mRNA linear VRT 12-NOV-2001  
LOCUS Gallus gallus adipocyte fatty acid binding protein (AFABP) mRNA,  
DEFINITION complete cds.  
ACCESSION AF432506  
VERSION AF432506.1 GI:16904068  
KEYWORDS chicken.  
SOURCE Gallus gallus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 472)  
AUTHORS Wang, Q., Li, N. and Li, H.  
TITLE Cloning and sequencing of adipocyte fatty acid binding protein gene in chicken

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 472)  
AUTHORS Wang, Q., Li, N. and Li, H.  
TITLE Direct Submission  
JOURNAL Submitted (12-OCT-2001) Animal Science and Technology College, Northeast Agricultural University, Mucai Street 59, Harbin, Heilongjiang 150030, China  
FEATURES Location/Qualifiers  
source 1..472  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
gene /tissue\_type="fat"  
CDS /dev\_stage="adult"  
/note="breed: broiler"  
1..472  
/gene="AFABP"  
35..433  
/gene="AFABP"  
/codon\_start=1  
/product="adipocyte fatty acid binding protein"  
/protein\_id="AAL30743.1"  
/db\_xref="GI:16904069"  
/translations="MCDQFVGTWKLSENFEYMKELGVGFATRKMGVAKPNLTIS INGDIVTIRSESTFKNTEISFKLGEFDETTADDRKTKNVITLDSGLKQVQKWDGKE TVIKRKVVVDGNLLVECTMNNVTSKRVYERA"  
BASE COUNT 153 a 85 c 131 g 103 t  
ORIGIN

Query Match 37.9%; Score 240; DB 5; Length 472;  
Best Local Similarity 72.2%; Pred. No. 1.8e-51;  
Matches 312; Conservative 0; Mismatches 120; Indels 0; Gaps 0;  
QY 55 ctcaacaagtgtgtagctttttagtagtacctggaactgtctccagtgaaaactttg 114  
Db 27 CTGCAAAATGTCGCCACCAAGTTTGTGGCACCTCGTGAAGCTCTTTCTAGTGAACCTTTG 86  
QY 115 atgattatataaagaagtagagtgaggctttgcccacagagaagagctgcgcagatgcca 174  
Db 87 AGGACTATATGAAGAGCTGGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 146  
QY 175 aacctaaacatgatcatcagtggaatgggagtgatgatacaccattataatctgaaagtacct 234  
Db 147 AGCCTAATTAATCATAGCATCAATGTGTGATGTGATACCATCATGATCAGAAAGTACCT 206  
QY 235 ttaaaaactgtagattctcttcatactctggtggcaggaaatttgacgaagtcaactgcagatg 294  
Db 207 TCAAAAATACAGAGATCTCTTTCAAGCTGGGTGAAGAGTTTGTATGAGACACACGAGATG 266  
QY 295 acaggaagagcaagagcaccataaccttagatgggggtgctcctgtacatgtgcagaaat 354  
Db 267 ACAGAAAAACAAAGAAATGTCAATACCCCTAGACAGTGGCACACTGAAGCAGGTGCAGAGT 326  
QY 355 gggatggaaaaatcaaacaccataaagagaaaaacagagagatgataaactggttggaat 414  
Db 327 GGGATGGCAAGAGACTGTTATCAAGAGAAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 386  
QY 415 gcgtcatgaaaggcgtcactctccacgagagatttatgagagatcaagaacgaaggacgtt 474  
Db 387 GCACCAATGAATAATCTTACCAGAAAAAGAGTTTACGAAAAAGACGATGAGGAAGCCGCTCTTC 446  
QY 475 gacctggactga 486  
Db 447 ATGTGGGACTGA 458  
RESULT 15  
RABPLP2  
LOCUS RABPLP2  
DEFINITION Rabbit myelin P2 mRNA, complete cds.  
ACCESSION J03744  
VERSION J03744.1 GI:165657  
KEYWORDS myelin.

SOURCE Rabbit (New Zealand white, 10 day old) sciatic nerve, cDNA to mRNA,  
clones pSN[2.1,2.2-2].

ORGANISM Oryctolagus cuniculus

REFERENCE 1 (bases 1 to 1836)  
AUTHORS Narayanan,V., Barbosa,E., Reed,R. and Tennekoon,G.  
TITLE Characterization of a cloned cDNA encoding rabbit myelin P2 protein.  
J. Biol. Chem. 263, 8332-8337 (1988)  
MEDLINE 88228063  
COMMENT Draft entry and printed copy of sequence for [1] kindly provided by  
G.Tennekoon, 11-MAR-1988.

FEATURES  
source  
1..1836  
/organism="Oryctolagus cuniculus"  
/db\_xref="taxon:9988"  
<1..1836  
46..444  
/product="myelin P2 mRNA"  
/note="myelin P2"  
/codon\_start=1  
/protein\_id="AAA31451.1"  
/db\_xref="GI:165658"  
/translation="MSNKLFGTWKLVSSNFDDYMKALGVGLATRKLGKLNKPNVVIS  
KKGDIITIRTESTFKNTEISFKLGQEFETADNRKTKSITTLERGALNQVQKWDGKE  
TTIKRLVDGKMVVECKMGVCTRIYEVK"

BASE COUNT 524 a 372 c 359 g 581 t  
ORIGIN Unreported.

Query Match 36.6%; Score 232; DB 4; Length 1836;  
Best Local Similarity 71.1%; Pred. No. 2.le-49;  
Matches 307; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

Qy 59 caaaatgtgtatccttttagtagtaacctggaactgtctccagtgaaacttttgatga 118  
Db 42 CACGATGAGCAACAAATTTCTAGGCACCTGGAAAGCTTGCTCCACGCAAAATTTTGATGA 101  
Qy 119 ttatatgaagaagtagtagtggtgtttgccaccaggaagtggtgscatggccaaacc 178  
Db 102 TTACATGAAGCTCTGGGTGTGGGTGTAGCCACCAGAAACTGGGAAATTTGGCCARACC 161  
Qy 179 taacatgatcatcagtggaatgggtgtgtatcaccatttaaatctgaaagtacatttaa 238  
Db 162 CAATGTGATCATCAGCAAGAAAGGGGATATCATCTATACGAACTGAAAGTACCTTCAA 221  
Qy 239 aaatactgagatttctctcattcctgagcaggaatttgacgaagtcactgcagatgacag 298  
Db 222 AATACTGAGATCTCCTTCAAGCTAGGCCAGGAATTTGAAGAAACCCACAGCTGACATAG 281  
Qy 299 gaaagtcaagagcaccataaccttagatgggggtgtcctgtacatgtgcagaaatggga 358  
Db 282 GAAACCAAGAGTATCATACCTTGGAAAGAGGCGCATTTGAATCAGGTACAGAAATGGGA 341  
Qy 359 tggaaatcaaccaccataaagagaacagagagatgataaacttggttggaatgagt 418  
Db 342 TGGCAAAAGAGACACCACTAAAGAGGAGTGTGGTGGGAAATGGTTGTGGAATGTAA 401  
Qy 419 catgaaggcgtcactccacgagagtttatgagagacataagcgaaggacgttgacc 478  
Db 402 GATGAAGGGCGTGTCTGCACCAGAACTCTATGAGAAGGCTCTGAGTAATTTTCTTCCTTGC 461  
Qy 479 tggactgaagtt 490  
Db 462 TGAAGTGATATT 473

Search completed: June 8, 2002, 08:43:34  
Job time: 9917 sec



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	651	96.4	132	11	Q99P60	Q99p60	spermophilu
2	603	89.3	150	11	Q9R290	Q9r290	rattus norv
3	537	79.6	113	6	Q9XSG4	Q9xsg4	oryctolagus
4	531	78.7	132	13	Q9OX56	Q9ox56	gallus gall
5	473.5	70.1	134	11	Q9QY04	Q9qy04	rattus norv
6	468	69.3	133	11	Q99P61	Q99p61	spermophilu
7	466	69.0	133	11	Q91W23	Q91w23	mus musculus
8	462	68.4	111	13	Q9OX55	Q9ox55	gallus gall
9	443	65.6	132	13	Q9OW92	Q9ow92	fundulus he
10	433	64.1	99	6	Q97675	Q97675	sus scrofa
11	427	63.3	125	6	Q9XS15	Q9xsi5	equus cabal
12	413	61.2	132	13	Q91BN9	Q91bn9	brachydantio
13	402	59.6	132	11	Q9DAL2	Q9dal2	mus musculus
14	400	59.3	133	13	Q97668	Q97668	cryodracon
15	397	58.8	137	13	Q97667	Q97667	chaenocepha
16	397	58.8	133	13	Q57670	Q57670	gobionototh







RESULT	13
Q9DAL2	
ID	PRELIMINARY; PRT; 132 AA.
Q9DAL2	
AC	Q9DAL2;
DT	01-JUN-2001 (TEMBLrel. 17, Created)
DT	01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT	01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE	1700007P1ORIK PROTEIN.
GN	1700007P1ORIK.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_Taxid=10090;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6J; TISSUE=TESTIS;
RC	MEDLINE=21085660; Pubmed=11217851;
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

GN	Cryodora antarcticus (Crocodylle icefish).
OS	H6-rABP.
ON	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Actinopterygii; Acanthopterygii; Percomorpha; Perciformes;
OC	Actinopterygii; Acanthopterygii; Percomorpha; Perciformes;
OC	Neoteleostei; Channichthyidae; Cryodora.
OX	NCBI_TaxID=36192;
OX	
OX	[1]
RP	SEQUENCE FROM N.A.
RP	TISSUE=HEART VENTRICLE;
RP	
RX	MEDLINE=98129752; PubMed=9461533;
RX	AYda M.E., Londraville R.L., Cashon R.E., Costello L., Sidell B.D.;
RT	"Two distinct types of fatty acid-binding protein are expressed in
RT	heart ventricle of Antarctic teleost fishes.";
RT	Biochem. J. 330:375-382(1998).
RL	EMBL; U92449; AAC60357.1; -;
RL	HSSP; P10790; 1BWY.
DR	InterPro; IPR000463; Fatty_acid_BP.
DR	

[illegible]

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 18:19:49 ; Search time 11.86 Seconds  
(without alignments)  
430.943 Million cell updates/sec

Title: US-09-503-596-4  
Perfect score: 675  
Sequence: 1 MCDAPGVGWLKLVSSSENFDDY.....KLVVECVMKGVSTRVYERA 132

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	670	99.3	131	1	FABA_HUMAN
2	620	91.9	131	1	FABA_MOUSE
3	608	90.1	131	1	FABA_PIG
4	607	89.9	131	1	FABA_RAT
5	573	84.9	131	1	FABA_BOVIN
6	482	71.4	131	1	MYP2_RABIT
7	472	69.9	132	1	FABH_RAT
8	466	69.0	132	1	FABH_MOUSE
9	459	68.0	131	1	MYP2_HUMAN
10	448	66.4	131	1	MYP2_BOVIN
11	446	66.1	131	1	MYP2_MOUSE
12	445	65.9	132	1	FABH_HUMAN
13	442	65.5	132	1	FABH_BOVIN
14	442	65.5	132	1	FABH_PIG
15	425	63.0	131	1	FABB_BOVIN
16	424	62.8	132	1	FABH_ONCMY
17	395	58.5	131	1	FABB_HUMAN
18	390	57.8	132	1	TLBP_RAT
19	388	57.5	131	1	FABB_CHICK
20	386	57.2	132	1	TLBP_MOUSE
21	381	56.4	131	1	FABB_RAT
22	378	56.0	132	1	FABL_GINGI
23	377	55.9	131	1	FABB_MOUSE
24	367	54.4	135	1	FABE_BOVIN
25	366	54.2	135	1	FABE_HUMAN
26	362	53.6	135	1	FABE_MOUSE
27	361	53.5	135	1	FABE_RAT
28	326	48.3	114	1	FABL_LAMJA
29	270.5	40.1	133	1	FABP_ECHGR
30	250	37.0	131	1	FABP_LEPDS
31	238	35.3	130	1	FABP_BLOTA
32	229.5	34.0	133	1	FABP_SCHMA
33	227	33.6	133	1	FABM_SCHGR

34	227	33.6	133	1	RET2_MOUSE	Q08652 mus musculu
35	225	33.3	133	1	RET2_HUMAN	P50120 homo sapien
36	225	33.3	133	1	RET2_RAT	P06768 rattus norv
37	223	33.0	133	1	RET2_PIG	P50121 sus scrofa
38	222	32.9	136	1	RET3_HUMAN	P29762 homo sapien
39	221	32.7	133	1	FABM_LOCM1	P41509 locusta mig
40	220	32.6	136	1	RET3_MOUSE	P02693 mus musculu
41	218	32.3	134	1	RET1_RAT	P02696 rattus norv
42	215	31.9	136	1	RET3_FUGRU	O42386 fugu rubrip
43	212	31.4	134	1	RET1_HUMAN	P09455 homo sapien
44	212	31.4	134	1	RET1_MOUSE	Q00915 mus musculu
45	206.5	30.6	137	1	RET4_MOUSE	P22935 mus musculu

ALIGNMENTS

RESULT 1  
FABA\_HUMAN  
ID FABA\_HUMAN STANDARD; PRT; 131 AA.  
AC P15090;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding  
DE protein) (ALBP) (A-FABP).  
GN FABP4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90105397; PubMed=2481498;  
RA Baxa C.A., Sha R.S., Buelt M.K., Smith A.J., Matarese V.,  
RA Chinander L.L., Boundy K.L., Bernlohr D.A.;  
RT "Human adipocyte lipid-binding protein: purification of the protein  
RT and cloning of its complementary DNA.";  
RL Biochemistry 28:8683-8690(1989).  
CC -!- FUNCTION: LIPID TRANSPORT PROTEIN IN ADIPOCYTES. BINDS BOTH LONG  
CC -!- CHAIN FATTY ACID AND RETINOIC ACID.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -!- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF  
CC TRANSPORTERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: J02874; AAA51689.1; -.  
DR PIR: A33363; A33363.  
DR HSP: P04117; 1L1D.  
DR MIM: 600434; -.  
DR InterPro: IPR000463; Fatty\_acid\_BP.  
DR InterPro: IPR000566; Lipocln\_cytFABP.  
DR Pfam: PF00061; lipocalin; 1.  
DR PRINTS: PR00178; FATTYACIDBP.  
DR PROSITE: PS00214; FABP; 1.  
DR Transprot: Lipid-binding; Phosphorylation.  
KW Transprot: Lipid-binding; Phosphorylation.  
FT INIT\_MET 0 BY SIMILARITY.  
FT MOD\_RES 19 19 PHOSPHORYLATION (BY TYR-KINASES)  
FT (BY SIMILARITY).  
SQ SEQUENCE 131 AA; 14588 MW; 81C241B63DD98235 CRC64;

Query Match 99.3%; Score 670; DB 1; Length 131;  
Best Local Similarity 100.0%; Pred. No. 1.4e-53;  
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```

FT STRAND 59 64
FT TURN 66 67
FT STRAND 70 73
FT TURN 75 76
FT STRAND 79 87
FT TURN 88 89
FT STRAND 90 97
FT TURN 98 99
FT STRAND 100 109
FT TURN 110 111
FT STRAND 112 119
FT TURN 120 121
FT STRAND 122 130
FT SEQUENCE 131 AA; 14519 MW; ED57D4E2774B8E32 CRC64;

Query Match 91.9%; Score 620; DB 1; Length 131;
Best Local Similarity 91.6%; Pred. No. 4.3e-49;
Matches 120; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

2 CDAFVGWTKLVSSNFDDYMKVGVGATRKVAGMKNMIIISVNGDVITIKSESTFKNT 61
|||||
1 CDAFVGWTKLVSSNFDDYMKVGVGATRKVAGMKNMIIISVNGDLVITIRSESTFKNT 60
|||||

QY 62 EISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKRRDDKLVVECVMK 121
|||||
DB 61 EISFKLGQEFDEVTADDRKVKSTITLDGGALVQVQKWDGKSTTIKRRDGDGLVVECVMK 120
|||||

QY 122 GVTSTRYVERA 132
|||||
DB 121 GVTSTRYVERA 131
|||||

RESULT 3
FABA_PIG STANDARD; PRT; 131 AA.
AC O97788;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding
DE protein) (ALBP) (A-FABP) (AP2).
GN FABP4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RN STRAIN=DUCOC; TISSUE=Liver;
RN MEDLINE=9909248; PubMed=9880671;
RA Gerbens F., Jansen A., van Erp A.J.M., Harders F., Meuwissen T.H.E.,
RA Rettenberger G.F.W., Veerkamp J.H., te Pas M.F.W.;
RT "The adipocyte fatty acid-binding protein locus: characterization and
RT association with intramuscular fat content in pigs.";
RL Mamm. Genome 9:1022-1026(1998).
RN [2]
RN SEQUENCE OF 36-47, AND TISSUE SPECIFICITY.
RP TISSUE=Fat;
RX MEDLINE=90241143; PubMed=2334399;
RA Armstrong M.K., Bernlohr D.A., Storch J., Clarke S.D.;
RT "The purification and characterization of a fatty acid binding protein
RT specific to pig (Sus domesticus) adipose tissue.";
RL Biochem. J. 267:373-378(1990).
CC -1- FUNCTION: LIPID TRANSPORT PROTEIN IN ADIPOCYTES. BINDS BOTH LONG
CC CHAIN FATTY ACID AND RETINOIC ACID.
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF INTRAMUSCULAR FAT
CC ACCRETION.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- TISSUE SPECIFICITY: ADIPOSE TISSUE.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y16039; CAA75995.1; -.
CC HSP; P04117; ILID.
CC InterPro: IPR000463; Fatty acid_BP.
CC InterPro: IPR000566; Lipocalin_cytFABP.
CC Pfam: PF00061; lipocalin; 1.
CC PRINTS; PR00178; FATTYACIDBP.
CC PROSITE; PS00214; FABP; 1.
KW Transport; Lipid-binding; Phosphorylation.
FT INIT_MET 0 BY SIMILARITY.
FT MOD_RES 19 19 PHOSPHORYLATION (BY TYR-KINASES)
FT SEQUENCE 131 AA; 14545 MW; 656CB0CA08CD4B2 CRC64;

Query Match 90.1%; Score 608; DB 1; Length 131;
Best Local Similarity 88.5%; Pred. No. 5.1e-48;
Matches 116; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 2 CDAFVGWTKLVSSNFDDYMKVGVGATRKVAGMKNMIIISVNGDVITIKSESTFKNT 61
|||||
DB 1 CDAFVGWTKLVSSNFDDYMKVGVGATRKVAGMKNLIIITVNGDMITIRSESTFKNT 60
|||||

QY 62 EISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKRRDDKLVVECVMK 121
|||||
DB 61 EIAFKLGQEFDEVTADDRKVKSTITLDGGALVQVQKWDGKTTTINRKIVDDKLVVECIMK 120
|||||

QY 122 GVTSTRYVERA 132
|||||
DB 121 GVTSTRYVERA 131
|||||

RESULT 4
FABA_RAT STANDARD; PRT; 131 AA.
AC P70623;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding
DE protein) (ALBP).
GN FABP4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RA Prinsen C., Veerkamp J.H.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: LIPID TRANSPORT PROTEIN IN ADIPOCYTES. BINDS BOTH LONG
CC CHAIN FATTY ACID AND RETINOIC ACID.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF
CC TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75581; AAB18344.1; -.

```

DR	PRINTS; PR00178; FATTYACIDBP.
DR	PROSITE; PS00214; FABP; 1.
KW	Transport; Lipid-binding; Phosphorylation.
FT	INIT_MET 0 BY SIMILARITY.
FT	MOD_RES 19 19 PHOSPHORYLATION (BY TYR-KINASES)
FT	(BY SIMILARITY).
SQ	SEQUENCE 131 AA; 14547 MW; A7D81B036C972D68 CRC64;

  

Query Match		84.9%;	Score 573;	DB 1;	Length 131;
Best Local Similarity		84.0%;	Pred.No. 7.le-45;		
Matches 110; Conservative		8;	Mismatches 13;	Indels	0; Gaps

  

QY	2	CDAFVGTVGKLVSSNFDDYMKVGVGFATRKVAGMAKPNNIIISVNGDVITIKSESTFKNT 61
		:     :     :     :     :     :     :     :     :     :
Db	1	CDAFVGTVGKLVSSNFDDYMKVGVGFATRKVAGMAKPTLIISLNGGVVTIKSESTFKNT 60
		:     :     :     :     :     :     :     :     :
QY	62	EISFILQGDFEVDATDRKVKSITILDGGVLHVHQWDGKSTTIKKRRDDKLVCVVK 121
		:     :     :     :     :     :     :     :     :
Db	61	EISFLKGDFEITPDORRVKYSIVNLDGALVQVNMDGKSTTIKKRLMDDRMVLECVMN 120
		:     :     :     :     :     :     :     :     :
QY	122	GVTSTRVYERA 132
		:
Db	121	GVTRATRYERA 131
		:

  

RESULT 6	
ID	MYP2_RABIT STANDARD; PRT; 131 AA.
AC	P02691;
DT	21-JUL-1986 (Rel. 01, Created)
DT	01-JUL-1989 (Rel. 11, Last sequence update)
DE	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Myelin P2 protein.
GN	PMP2.
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=88228063; PubMed=2453513;
RA	Narayanan V., Barbosa E., Reed R., Tennekoon G.;
RT	"Characterization of a cloned cDNA encoding rabbit myelin P2
RL	protein.";
RL	J. Biol. Chem. 263:8332-8337(1988).
RP	[2]
RP	SEQUENCE OF 1-55.
RX	MEDLINE=80094496; PubMed=7356651;
RA	Ishaque A., Hofmann T., Rhees S., Eylar E.H.;
RT	"The NH <sub>2</sub> -terminal region of the P2 protein from rabbit sciatic nerve
RL	myelin.";
RL	J. Biol. Chem. 255:1058-1063(1980).
RN	[3]
RP	SEQUENCE OF 55-131.
RX	MEDLINE=82098098; PubMed=6172423;
RA	Ishaque A., Hofmann T., Eylar E.H.;
RT	"The complete amino acid sequence of the rabbit P2 protein.";
RL	J. Biol. Chem. 257:592-595(1982).
CC	-I- FUNCTION: THIS PROTEIN MAY BE A LIPID TRANSPORT PROTEIN IN SCHWANN CELLS.
CC	-I- MISCELLANEOUS: P2 PROTEIN AND MYELIN BASIC PROTEIN TOGETHER CONSTITUTE A MAJOR FRACTION OF PERIPHERAL NERVOUS SYSTEM MYELIN PROTEIN.
CC	-I- SIMILARITY: BELONGS TO THE FABP/P2/CBPP/FAMILY OF TRANSPORTERS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collabor-
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstat-
CC	the European Bioinformatics Institute. There are no restrictions on
CC	use by non-profit institutions as long as its content is in no
CC	modified and this statement is not removed. Usage by and for commo-
CC	entities requires a license agreement (See http://www.isb-sib.ch/anno-

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC -----  
DR EMBL; J03744; AAA31451.1; -  
DR PIR; A03145; MPRB2.  
DR PIR; A28081; A28081.  
DR HSP; P02690; 1PMP.  
DR InterPro; IPR000463; Fatty\_acid\_BP.  
DR InterPro; IPR000586; Lipocin\_cycFABP.  
DR Pfam; PF00061; Lipocalin; 1.  
DR PRINTS; PR00178; FATTYACIDBP.  
DR PROSITE; PS00214; FABP; 1.  
KW Myelin; Lipid-binding; Transport; Acetylation.  
FT INIT\_MET 0 0  
FT MOD\_RES 0 1 ACETYLTATION.  
FT CONFLICT 72 72 E -> Q (IN REF. 3).  
FT CONFLICT 83 83 I -> T (IN REF. 3).  
FT CONFLICT 98 98 D -> N (IN REF. 3).  
SQ SEQUENCE 131 AA; 14791 MW; E53788F693BC6A0E CRC64;  
  
Query Match 71.4%; Score 482; DB 1; Length 131;  
Best Local Similarity 72.4%; Pred. No. 1e-36;  
Matches 92; Conservative 14; Mismatches 21; Indels 0; Gaps 0;  
  
QY 5 FVGTWKLVSSEDDYMKVGVGFATKRVAGMAKPNMIIISVNGDVITIKSETFKNTAIS 64  
DB 4 FLGTWKLVSSEDDYMKVLGVGLATKRLGNLAKPNVLIISKGGDIITRTSETFKNTAIS 63  
QY 65 FILQGEFDEVTADRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLWVECVKMGVT 124  
DB 64 FLQGEFEETADNRKTSIITLGRALNQVQKWDGKETIKRLVDGKVVVECKMGVV 123  
QY 125 STRVYER 131  
DB 124 CTRIYEK 130  
|:|:|:  
|:|:|:  
  
RESULT 7  
FABH\_RAT STANDARD; PRT; 132 AA.  
AC P07483;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DE 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Fatty acid-binding protein, heart (H-FABP).  
GN FABP3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87250640; PubMed=3036869;  
RA Heuckeroth R.O., Birkenmeier E.H., Levin M.S., Gordon J.I.;  
RT "Analysis of the tissue-specific expression, developmental  
RT regulation, and linkage relationships of a rodent gene encoding heart  
RT fatty acid binding protein.";  
RL J. Biol. Chem. 262:9709-9717(1987).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88107756; PubMed=3427112;  
RA Claffey K.P., Herrera V.L., Brecher P., Ruiz-Opazo N.;  
RT "Cloning and tissue distribution of rat heart fatty acid binding  
RT protein mRNA: identical forms in heart and skeletal muscle.";  
RL Biochemistry 26:7900-7904(1987).  
[3]  
RP SEQUENCE  
RX MEDLINE=88153733; PubMed=3162235;  
RA Gibson B.W., Yu Z., Aberth W., Burlingame A.L., Bass N.M.;  
RT "Revision of the blocked N terminus of rat heart fatty acid-binding  
RT protein by liquid secondary ion mass spectrometry.";  
RL J. Biol. Chem. 263:4182-4185(1988).  
[4]

RP PRELIMINARY SEQUENCE.  
RX MEDLINE=86250713; PubMed=2424895;  
RA Sacchetti J.C., Said B., Schulz H., Gordon J.I.;  
RT "Rat heart fatty acid-binding protein is highly homologous to the  
RT murine adipocyte 422 protein and the P2 protein of peripheral nerve  
RT myelin.";  
RL J. Biol. Chem. 261:8218-8223(1986).  
[5]  
RP SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=91170283; PubMed=2005132;  
RA Kimura H., Odani S., Nishi S.-I., Sato H., Arakawa M., Ono T.;  
RT "Primary structure and cellular distribution of two fatty  
RT acid-binding proteins in adult rat kidneys.";  
RL J. Biol. Chem. 266:5963-5972(1991).  
[6]  
RP SEQUENCE OF 58-86.  
RX MEDLINE=89374061; PubMed=2775193;  
RA Kimura H., Hitomi M., Odani S., Koide T., Arakawa M., Ono T.;  
RT "Rat heart fatty acid-binding protein. Evidence that supports the  
RT amino acid sequence predicted from the cDNA.";  
RL Biochem. J. 260:303-306(1989).  
[7]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Stomach;  
RX MEDLINE=90032682; PubMed=2806260;  
RA Kanda T., Iseki S., Hitomi M., Kimura H., Odani S., Kondo H.,  
RA Matsubara Y., Muto T., Ono T.;  
RT "Purification and characterization of a fatty-acid-binding protein  
RT from the gastric mucosa of rats. Possible identity with heart  
RT fatty-acid-binding protein and its parietal cell localization.";  
RL Eur. J. Biochem. 185:27-33(1989).  
[8]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=88326235; PubMed=3415652;  
RA Jones P.D., Carne A., Bass N.M., Grigor M.R.;  
RT "Isolation and characterization of fatty acid binding proteins from  
RT mammary tissue of lactating rats.";  
RL Biochem. J. 251:919-925(1988).  
[9]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Mammary gland;  
RX MEDLINE=94162301; PubMed=8117746;  
RA Nielsen S.O., Rump R., Hoejrup P., Roepstorff P., Spener F.;  
RT "Differential regulation and phosphorylation of the fatty acid-  
RT binding protein from rat mammary epithelial cells.";  
RL Biochim. Biophys. Acta 1211:189-197(1994).  
CC -|- FUNCTION: FABP ARE THOUGHT TO PLAY A ROLE IN THE INTRACELLULAR  
CC TRANSPORT OF LONG-CHAIN FATTY ACIDS AND THEIR ACYL-COA ESTERS.  
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -|- TISSUE SPECIFICITY: HEART, BUT ALSO SKELETAL MUSCLE, KIDNEY,  
CC BRAIN AND MAMMARY GLAND.  
CC -|- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRABP FAMILY OF  
CC TRANSPORTERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02773; AAA41136.1; -  
DR EMBL; M18034; AAA41137.1; -  
DR PIR; A23838; A23838.  
DR PIR; A27452; A27452.  
DR PIR; A28197; A28197.  
DR PIR; A28458; A28458.  
DR PIR; A39551; A39551.  
DR PIR; S06478; S06478.  
DR HSP; P05413; 1HMT.





```
RESULT 11
MYP2_MOUSE STANDARD; PRT; 131 AA.
ID P24526;
AC
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myelin P2 protein.
GN PMP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91268811; PubMed=1711100;
RA Narayanan V., Kaestner K.H., Tennekoon G.I.;
RT "Structure of the mouse myelin P2 protein gene.";
RL J. Neurochem. 57:75-80(1991).
CC -1- FUNCTION: THIS PROTEIN MAY BE A LIPID TRANSPORT PROTEIN IN SCHWANN
CC CELLS.
CC -1- MISCELLANEOUS: P2 FRACTION AND MYELIN BASIC PROTEIN TOGETHER
CC CONSTITUTE A MAJOR FRACTION OF PERIPHERAL NERVOUS SYSTEM MYELIN
CC PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE FABP/P2/CRBP/CRAPP FAMILY OF
CC TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S39508; AAB19249.2; -
DR PIR; JH0407; JH0407.
DR HSSP; P02690; 1PMP.
DR MGD; MGI:102667; Pmp2.
DR InterPro; IPR000463; Fatty_acid_BP.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR Pfam; PF00061; Lipocalin; 1.
DR PRINTS; PR00178; FATTYACIDBP.
DR PROSITE; PS00214; FABP; 1.
KW Myelin; Lipid-binding; Transport; Acetylation.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 131 AA; 14804 MW; AB357746F152E07C CRC64;

Query Match 66.18; Score 446; DB 1; Length 131;
Best Local Similarity 66.18; Pred. No. 1.7e-33;
Matches 84; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

Qy 5 FVGTWKLVSSEFDYMKYGVGFKTRKVGAKPMNIIISVNGDVITIKSESTFKNTEIS 64
I:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Dy 4 FLGTWKLVSSEHFDYMKALGVGLNKLGNLAKPTVIISKGDVITIRTESAFKNTEIS 63

Qy 65 FILGQEFDEVTDARRKVKSTTLTGGVLVHVQKWDGKSTTKRKREDKLKVECVKMGVT 124
I:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Dy 64 FKLGGFEDETTADNRKAKSVITLGRSLQVQKWDGKETAIRRTLLDGRMVVECIKMGW 123

Qy 125 STRVYER 131
I:||||:
Dy 124 CTRIYER 130

RESULT 12
FABH_HUMAN STANDARD; PRT; 132 AA.
ID P03413; Q99957;
AC
```

```
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Fatty acid-binding protein, heart (H-FABP) (Muscle fatty acid-binding
DE protein) (M-FABP) (Mammary-derived growth inhibitor) (MDGI).
GN FABP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX TISSUP=Skeletal muscle;
RX MEDLINE=91248148; PubMed=1710107;
RA Pieter R.A., Veerkamp J.H., Kanda T., Ono T., Geurts van Kessel A.;
RT "Cloning of the cDNA encoding human skeletal-muscle
RT fatty-acid-binding protein, its peptide sequence and chromosomal
RT localization.";
RL Biochem. J. 276:203-207(1991).
RN [2]
SEQUENCE FROM N.A.
RX TISSUP=Breast;
RA Hu Y.F., Ao X., Russo I.H., Russo J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Wu X., Arlt M., Goodfellow P.J., Rottman J.N.;
RT "Genomic organization and complete nucleotide sequence of the human
RT cardiac fatty acid binding protein gene (FABP3), and identification of
RT a closely related genomic sequence.";
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE.
RX MEDLINE=88339792; PubMed=3421901;
RA Offner G.D., Brecher P., Sawilivich W.B., Costello C.E., Troxler R.F.;
RT "Characterization and amino acid sequence of a fatty acid-binding
RT protein from human heart.";
RL Biochem. J. 252:191-198(1988).
RN [5]
REVIEWS, AND SEQUENCE.
RX MEDLINE=91094793; PubMed=2266954;
RA Boerchers T., Hoejrup P., Nielsen S.U., Roepstorff P., Spener F.,
RA Knudsen J.;
RT "Revision of the amino acid sequence of human heart fatty
RT acid-binding protein.";
RL Mol. Cell. Biochem. 98:127-133(1990).
RN [6]
SEQUENCE OF 14-132 FROM N.A.
RX TISSUP=Heart;
RX MEDLINE=94085953; PubMed=8262516;
RA Troxler R.F., Offner G.D., Jiang J.W., Wu B.L., Skare J.C.,
RA Milunsky A., Wyandt H.E.;
RT "Localization of the gene for human heart fatty acid binding protein
RT to chromosome 1p32-1p33.";
RL Hum. Genet. 92:563-566(1993).
RN [7]
SEQUENCE OF 31-38.
RX TISSUP=Heart;
RX MEDLINE=96007936; PubMed=7498159;
RA Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A.,
RA Ershova E.S., Egorov T.A., Musalyamov A.K.;
RT "The major protein expression profile and two-dimensional protein
RT database of human heart.";
RL Electrophoresis 16:1160-1169(1995).
RN [8]
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=92406763; PubMed=1526991;
RA Zanotti G., Scapin G., Spadon P., Veerkamp J.H., Sacchettini J.C.;
RT "Three-dimensional structure of recombinant human muscle fatty acid-
RT binding protein.";
RL J. Biol. Chem. 267:18541-18550(1992).
RN [9]
X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
```



"Purification and characterisation of a polymorphic low M(r) bovine muscle cysteine proteinase inhibitor: structural identity with fatty-acid-binding proteins.";  
Biochimie 75:937-945(1993).

[6]

VARIANT IN POSITION 98.

MEDLINE-9335876; PubMed-8354262;

Bartetzko N., Lezius A.G., Spener F.;

"Isoforms of fatty-acid-binding protein in bovine heart are coded by

distinct mRNA.";

Eur. J. Biochem. 215:555-559(1993).

[7]

X-RAY CRYSTALLOGRAPHY.

MEDLINE-91301134; PubMed-2070787;

Mueller-Fahrnow A., Egner U., Jones T.A., Ruedel H., Spener F.,

Saenger W.;

"Three-dimensional structure of fatty-acid-binding protein from

bovine heart.";

Eur. J. Biochem. 199:271-276(1991).

[8]

STRUCTURE BY NMR.

MEDLINE-93130916; PubMed-1483473;

Luecke C., Lassen D., Kreienkamp H.-J., Spener F., Rueterjans H.;

"Sequence-specific 1H-NMR assignment and determination of the

secondary structure of bovine heart fatty-acid-binding protein.";

Eur. J. Biochem. 210:901-910(1992).

[9]

STRUCTURE BY NMR.

MEDLINE-95324534; PubMed-7601110;

Lassen D., Luecke C., Kveder M., Mesgarzadeh A., Schmidt J.M.,

Specht B., Lezius A., Spener F., Rueterjans H.;

"Three-dimensional structure of bovine heart fatty-acid-binding

protein with bound palmitic acid, determined by multidimensional NMR

spectroscopy.";

Eur. J. Biochem. 230:266-280(1995).

[10]

STRUCTURE BY NMR.

MEDLINE-98149696; PubMed-9490052;

Mesgarzadeh A., Pfeiffer S., Engelke J., Lassen D., Rueterjans H.;

"Bound water in apo and holo bovine heart fatty-acid-binding protein

determined by heteronuclear NMR spectroscopy.";

Eur. J. Biochem. 251:781-786(1998)

-1- FUNCTION: FAPB ARE THOUGHT TO PLAY A ROLE IN THE INTRACELLULAR

TRANSPORT OF LONG-CHAIN FATTY ACIDS AND THEIR ACYL-COA ESTERS.

-1- FUNCTION: MDGI REVERSELY INHIBITS PROLIFERATION OF MAMMARY

CARCINOMA CELLS.

-1- SUBCELLULAR LOCATION: CYTOPLASMIC OR MITOCHONDRIAL MATRIX.

-1- TISSUE SPECIFICITY: MAMMARY EPITHELIAL CELLS OF DEVELOPING

LOBULOALVEOLAR STRUCTURES AND HEART.

-1- SIMILARITY: BELONGS TO THE FAPB/P2/CRBP/CRABP FAMILY OF

TRANSPORTERS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X12710; CAA31212.1; -;

DR EMBL; X51933; CAA36199.1; -;

DR PIR; A29466; A29466.

DR PIR; S01133; S01133.

DR PIR; A34676; A34676.

DR PIR; A37883; A37883.

DR PDB; 1BWY; 07-OCT-98.

DR InterPro; IPR000463; Fatty\_acid\_BP.

DR InterPro; IPR000566; Lipocln\_cytFAPB.

DR Pfam; PF00061; Lipocalin; 1.

DR PRINTS; PR00178; FATTYACIDBP.

DR PROSITE; PS00214; FAPB; 1.

DR Transport; Lipid-binding; Acetylation; Phosphorylation; 3D-structure.

KW

FT INIT\_MET 0 0  
FT MOD\_RES 1 1  
FT MOD\_RES 19 19  
FT MOD\_RES 19 19  
FT VARIANT 98 98  
FT CONFLICT 12 14  
FT CONFLICT 40 40  
FT CONFLICT 43 43  
FT CONFLICT 93 93  
FT CONFLICT 127 127  
SQ SEQUENCE 132 AA; 14648 MW; C6204932C6DFD35 CRC64;

Query Match 65.5%; Score 442; DB 1; Length 132;  
Best Local Similarity 65.1%; Pred. No. 4e-33;  
Matches 84; Conservative 15; Mismatches 30; Indels 0; Gaps 0;

Qy 3 DAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIIISVNGDVITIKSESTFKNTE 62

Db 2 DAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIIISVNGDVITIKSESTFKNTE 61

Qy 63 ISFTILGQEFDEVTADDRKVKSTITLDGCVLHVHVKWCKSTIKRREDDKLIVCEVMKG 122

Db 62 ISFTILGQEFDEVTADDRKVKSTITLDGCVLHVHVKWCKSTIKRREDDKLIVCEVMKG 121

Qy 123 VTSTRVYER 131

Db 122 AVCRTYEK 130

RESULT 14

FABH\_PIG

ID FABH\_PIG STANDARD; PRG; 132 AA.

AC 002772;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Fatty acid-binding protein, heart (H-FABP).

GN FAPB3.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=97263098; PubMed=9107676;

RA Gerbens F.N.A., Rettenberger G., Lenstra J.A., Veerkamp J.H.,

Te Pas M.F.W.;

RT "Characterization, chromosomal localization, and genetic variation of

the porcine heart fatty acid-binding protein gene.";

RL Mamm. Genome 8:328-332(1997).

CC -1- FUNCTION: FAPB ARE THOUGHT TO PLAY A ROLE IN THE INTRACELLULAR

TRANSPORT OF LONG-CHAIN FATTY ACIDS AND THEIR ACYL-COA ESTERS.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO THE FAPB/P2/CRBP/CRABP FAMILY OF

TRANSPORTERS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X98558; CAA67168.1; -;

DR HSSP; P10790; IEWY.

DR InterPro; IPR000463; Fatty\_acid\_BP.

DR InterPro; IPR000566; Lipocln\_cytFAPB.

DR Pfam; PF00061; Lipocalin; 1.

DR PRINTS; PR00178; FATTYACIDBP.

DR PROSITE; PS00214; FAPB; 1.

DR PROSITE; PS00214; FAPB; 1.



Search completed: June 7, 2002, 18:23:20  
Job time: 211 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 18:17:38 ; Search time 16.41 Seconds  
(without alignments)  
772.931 Million cell updates/sec

Title: US-09-503-596-4  
Perfect score: 675  
Sequence: 1 MCDAFVGWKLVSSENFDDY.....KLVECVMGVSTRVYERA 132

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	675	100.0	132	1 FZHUF	fatty acid-binding
2	623	92.3	132	2 B25952	myelin P2 protein
3	578	85.6	132	2 S57744	adipocyte-type fat
4	486	72.0	132	1 MPRB2	myelin P2 protein
5	477	70.7	133	2 A27452	fatty acid-binding
6	466	69.0	133	2 PC4011	fatty acid-binding
7	463	68.6	132	1 MPRU2	myelin P2 protein
8	450	66.7	132	2 JH0407	myelin P2 protein
9	449	66.5	133	1 F2HUC	fatty acid-binding
10	448	66.4	131	1 MPB02	myelin P2 protein
11	446	66.1	133	2 A34676	fatty acid-binding
12	425	63.0	131	2 S06479	fatty acid-binding
13	392	58.1	132	2 A49184	fatty acid-binding
14	390	57.8	132	2 I52524	testis lipid bindi
15	385	57.0	132	2 I56510	fatty acid-binding
16	381	56.4	132	2 I58161	lipid-binding prot
17	381	56.4	132	2 I48923	fatty acid-binding
18	378	56.0	132	2 S20297	fatty acid-binding
19	366	54.2	135	2 I56326	fatty acid-binding
20	362	53.6	135	2 A47497	lipid-binding prot
21	361	53.5	135	2 JC2201	fatty acid-binding
22	307.5	45.6	135	2 T15205	hypothetical prote
23	300.5	44.5	136	2 T15207	hypothetical prote
24	286.5	38.0	137	2 T25123	hypothetical prote
25	242.5	35.9	133	2 A48578	fatty acid-binding
26	231.5	34.3	134	2 S29600	fatty acid-binding
27	229.5	34.0	133	2 A39818	14K fatty acid-bin
28	227	33.6	133	2 A44870	fatty acid-binding
29	227	33.6	134	2 S34717	retinol-binding pr

30	225	33.3	134	2 A29065	retinol-binding pr
31	223	33.0	134	2 S45378	cellular retinol b
32	223	33.0	134	2 S43470	fatty acid-binding
33	222	32.9	137	1 RJHU1	retinoic acid-bind
34	220	32.6	137	1 RJBOA	retinoic acid-bind
35	220	32.6	137	2 A35825	retinoic acid-bind
36	218	32.3	135	1 RJRTO	retinol-binding pr
37	212	31.4	135	1 RJHU0	retinol-binding pr
38	212	31.4	135	2 S16355	retinol-binding pr
39	206	30.5	134	2 S69360	retinoic acid-bind
40	203.5	30.1	138	1 A42495	retinoic acid-bind
41	201.5	29.9	138	1 RJHU2	retinoic acid-bind
42	197.5	29.3	138	2 I51265	xCRABP - African c
43	196	29.0	139	2 I53298	cellular retinoic
44	191	28.3	95	2 A61629	retinoic acid-bind
45	190	28.1	132	1 FZHUI	fatty acid-binding

ALIGNMENTS

RESULT 1

FZHUF  
fatty acid-binding protein, adipocyte - human  
N:Alternate names: adipocyte lipid-binding protein (ALBP); fatty acid binding protein  
C:Species: Homo sapiens (man)  
C:Date: 20-Dec-1989 #sequence\_revision 03-May-1996 #text\_change 22-Jun-1999  
C:Accession: A33363  
R:Baxa, C.A.; Sha, R.S.; Buelt, M.K.; Smith, A.J.; Matarese, V.; Chinander, L.L.; Bou  
Biochemistry 28, 8683-8690, 1989  
A:Title: Human adipocyte lipid-binding protein: purification of the protein and cloni  
A:Reference number: A33363; MUID: 90105297  
A:Accession: A33363  
A:Molecule type: mRNA  
A:Residues: 1-132 <BAX>  
A:Cross-references: GB:J02874; NID:gl78346; PIDN:AAA51689.1; PID:gl78347  
A:Experimental source: subcutaneous fat  
C:Genetics:  
A:Gene: GDB:FABP4  
A:Cross-references: GDB:128030  
A:Map position: lp33-1p32  
C:Superfamily: myelin P2 protein  
C:Keywords: blocked amino end; lipid binding; phosphoprotein  
F:2/Modified site: blocked amino end (Cys) (in mature form) (probably acetylated) #st  
F:20/Binding site: phosphate (Tyr) (covalent) #status predicted  
F:127,129/Binding site: fatty acid (Arg, Tyr) #status predicted

Query Match 100.0%; Score 675; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 8.3e-55;  
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCDAFVGWKLVSSENFDDYMKVEGVGFATRKVAGMAKPMIISVNGDVITIKSESTFKN 60

Db 1 MCDAFVGWKLVSSENFDDYMKVEGVGFATRKVAGMAKPMIISVNGDVITIKSESTFKN 60

Qy 61 TEISILGQEFDEVADRRKVKSTITLDGVLVHVQKWDGKSTTTIKRREDKLVVEVCYM 120

Db 61 TEISILGQEFDEVADRRKVKSTITLDGVLVHVQKWDGKSTTTIKRREDKLVVEVCYM 120

Qy 121 KGVSTRVYERA 132

Db 121 KGVSTRVYERA 132

RESULT 2

B25952

myelin P2 protein homolog - mouse

N:Alternate names: adipocyte lipid-binding protein; adipocyte P2 protein

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 31-Dec-2000

C:Accession: B25952; PC1249; A24884; A05089; A30810; I49440; I49467

R:Phillips, M.; Djian, P.; Green, H.

J. Biol. Chem. 261, 10821-10827, 1986

A:Title: The nucleotide sequence of three genes participating in the adipose differentiation

A:Reference number: A92553; MUID:86278164

A:Accession: B25952

A:Molecule type: DNA

A:Residues: 1-132 <PHI>

A:Cross-references: GB:M13385; NID:gl98718; PIDN:AAA39417.1; PID:g387401

A:Note: the authors translated the codon GTG for residue 111 as Gly

R:Bansal, M.P.; Medina, D.

Biochem. Biophys. Res. Commun. 191, 61-69, 1993

A:Title: Expression of fatty acid-binding proteins in the developing mouse mammary gland

A:Reference number: PC1248; MUID:93191717

A:Accession: PC1249

A:Molecule type: protein

A:Residues: 25-35;37-51;59-88 <BAN>

A:Experimental source: mammary gland

R:Hunt, C.R.; Ro, J.H.S.; Dobson, D.E.; Min, H.Y.; Spiegelman, B.M.

Proc. Natl. Acad. Sci. U.S.A. 83, 3786-3790, 1986

A:Title: Adipocyte P2 gene: Developmental expression and homology of 5'-flanking sequence

A:Reference number: A24884; MUID:86233319

A:Accession: A24884

A:Molecule type: DNA

A:Residues: 1-39,'T','A','110','G','112-132 <HUN>

A:Cross-references: GB:M13264; NID:g200183; PIDN:AAA39870.1; PID:g387504

A:Note: the authors translated the codon ACC for residue 40 as Ser

R:Bernlohr, D.A.; Angus, C.W.; Lane, M.D.; Bolanowski, M.A.; Kelly Jr., T.J.

Proc. Natl. Acad. Sci. U.S.A. 81, 5468-5472, 1984

A:Title: Expression of specific mRNAs during adipose differentiation: identification of

A:Reference number: A05089; MUID:84298159

A:Accession: A05089

A:Molecule type: mRNA

A:Residues: 1-110,'G','112-132 <BER>

A:Cross-references: GB:X02109; NID:gl98716; PIDN:AAA39416.1; PID:g293695

R:Matarse, V.; Bernlohr, D.A.

J. Biol. Chem. 263, 14544-14551, 1988

A:Title: Purification of murine adipocyte lipid-binding protein. Characterization as a

A:Reference number: A30810; MUID:89008309

A:Accession: A30810

A:Molecule type: protein

A:Residues: 2-110,'G','112-132 <MAT>

R:Cook, K.S.; Hunt, C.R.; Spiegelman, B.M.

J. Cell Biol. 100, 514-520, 1985

A:Title: Developmentally regulated mRNAs in 3T3-adipocytes: Analysis of transcriptional

A:Reference number: I49440; MUID:85105214

A:Accession: I49440

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 11-110,'G','112-132 <RES>

A:Cross-references: GB:M28726; NID:gl91492; PIDN:AAA37112.1; PID:gl91493

R:Cook, J.S.; Lucas, J.J.; Sibley, E.; Bolanowski, M.A.; Christy, R.J.; Kelly, T.J.; Lan

Proc. Natl. Acad. Sci. U.S.A. 85, 2949-2953, 1988

A:Title: Expression of the differentiation-induced gene for fatty acid-binding protein

A:Reference number: I49467; MUID:88203618

A:Accession: I49467

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-11 <RE2>

A:Cross-references: GB:M20497; NID:gl91743; PIDN:AAA37188.1; PID:gl91744

A:Experimental source: 3T3-L1 cells

C:Genetics:

C:Superfamily: myelin P2 protein

C:Introns: 25/1; 82/3; 116/3

C:Keywords: lipid binding; phosphoprotein

Query Match	92.3%	Score 623;	DB 2;	Length 132;
Best Local Similarity	91.7%	Pred. No. 4.8e-50;		
Matches 121;	Conservative	5;	Mismatches 6;	Indels 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGDVTIIRSESTFKN 60

Db 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGDVTIIRSESTFKN 60

Qy	61	TEISFILGQEFDEVTADRRKVKSTTTLDGGVLVHVQKWDGKSTTIKKRREDDKLVVECV	120
		:       :       :       :       :       :       :       :       :	
Db	61	TEISFLGVFEDEITADRRKVKSTTTLDGGALVQVQKWDGKSTTIKKRROVDKLVVECV	120
Qy	121	KGVTSTRVYERA	132
		:       :       :       :       :       :       :       :       :	
Db	121	KGVTSTRVYERA	132

  

RESULT	3	
S57744		
adipocyte-type fatty acid binding protein - bovine		
C:Species: Bos primigenius taurus (cattle)		
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999		
C:Accession: S57744		
R:Specht, B.; Bartetzko, N.; Kuhl, H.; Franke, R.; Boerchers, T.; Spener, F.		
submitted to the EMBL Data Library, June 1995		
A:Description: Mammary derived growth inhibitor - not a distinct protein but		
A:Reference number: S57744		
A:Accession: S57744		
A>Status: preliminary		
A:Molecule type: mRNA		
A:Residues: 1-132 <SPE>		
A:Cross-references: EMBL:X89244; NID:g895753; PIDN:CAA61532.1; PID:g895754		
C:Superfamily: myelin P2 protein		

  

Query Match	85.6%	Score 578;	DB 2;	Length 132;	
Best Local Similarity	84.1%;	Pred. No. 6.4e-46;			
Matches 111;	Conservative	8;	Mismatches 13;	Indels 0;	Gaps

  

Qy	1	MCDAFVGTWKLVSSENFDDYMKVEVGVGFATRKVAGMAKPNMIIISVNGDVITIKSESTFKN	60
		:       :       :       :       :       :       :       :       :	
Db	1	MCDAFVGTWKLVSSENFDDYMKVEVGVGFATRKVAGMAKPTLIISLNGGVVTVTIKSESTFKN	60
Qy	61	TEISFILGQEFDEVTADRRKVKSTTTLDGGVLVHVQKWDGKSTTIKKRREDDKLVVECV	120
		:       :       :       :       :       :       :       :       :	
Db	61	TEISFLGQEFDEITPDRRKVKSIYNLDEGALVQVQKWDGKSTTIKKRLMDDRWLVVECV	120
Qy	121	KGVTSTRVYERA	132
		:       :       :       :       :       :       :       :       :	
Db	121	NGVTATRVYERA	132

  

RESULT	4	
MPRE2		
myelin P2 protein - rabbit		
C:Species: Oryctolagus cuniculus (domestic rabbit)		
C:Date: 30-Jun-1980 #sequence_revision 10-May-1996 #text_change 22-Jun-1999		
C:Accession: A28081, A92266; A92372; A03145		
R:Narayanan, V.; Barbosa, E.; Reed, R.; Tennekoon, G.		
J. Biol. Chem. 263, 8332-8337, 1988		
A:Title: Characterization of a cloned cDNA encoding rabbit myelin P-2 protei		

RESULT 4

MPRE2

myelin P2 protein - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Jun-1980 #sequence.revision 10-May-1996 #text\_change 22-Jun-1999

C:Accession: A28081; A92266; A92372; A03145

R:Narayanan, V.; Barbosa, E.; Reed, R.; Tennekoon, G.

J. Biol. Chem. 263, 8332-8337, 1988

A:Title: Characterization of a cloned cDNA encoding rabbit myelin P-2 protein.

A:Reference number: A28081; MUID:88228063

A:Accession: A28081

A:Molecule type: mRNA

A:Residues: 1-132 <NAR>

A:Cross-references: GB:J03744; NID:gl65657; PIDN:AAA31451.1; PID:gl65658

A:Note: translation of initiator Met is not shown

R:Ishaque, A.; Hofmann, T.; Rhee, S.; Eylar, E.H.

J. Biol. Chem. 255, 1058-1063, 1980

A:Title: The NH-2-terminal region of the P2 protein from rabbit sciatic nerve myelin.

A:Reference number: A92266; MUID:80094496

A:Accession: A92266

A:Molecule type: protein

A:Residues: 2-56 <IS1>

R:Ishaque, A.; Hofmann, T.; Eylar, E.H.

J. Biol. Chem. 257, 592-595, 1982

A:Title: The complete amino acid sequence of thr rabbit P2 protein.

A:Reference number: A92372; MUID:82098098

A:Accession: A92372

A:Molecule type: protein

A:Residues: 55-72, 'Q', 74-83, 'T', 85-98, 'N', 100-132 <IS2>

**A:Title:** Expression of fatty acid-binding proteins in the developing mouse mammary gland

A:Reference number: PC1248; MUID:93191717

A:Accession: PC1248

A:Molecule type: protein

A:Residues: 23-108,'R',110-119 <BAN>

A:Experimental source: mammary gland

C:Comment: This protein functions as an endogenous growth inhibitor and differentiation

C:Genetics:

A:Gene: mdg1/h-fabp

A:Introns: 24/1; 82/2; 133/1

C:Superfamily: myelin P2 protein

C:Keywords: lipid binding

Query Match 69.0%; Score 466; DB 2; Length 133;

Best Local Similarity 67.9%; Pred. No. 1.2e-35;

Matches 89; Conservative 15; Mismatches 27; Indels 0; Gaps 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGVDVIKSESTFKN 60

Db 1 MADAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGVDVIKSESTFKN 60

Qy 61 TEISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120

Db 61 TEINFGLGIEFDEVTADDRKVKSLVTLDDGGLIHKVQKNGQETTLTRELVDGKLILTLH 120

Qy 121 KGVTSRTRYER 131

Db 121 GSVVSTRTRYEK 131

RESULT 7

MPH02

N:myelin P2 protein [validated] - human

N:Alternate names: peripheral myelin protein 2

C:Species: Homo sapiens (man)

C:Date: 05-Apr-1983 #sequence\_revision 27-Jan-1995 #text\_change 08-Dec-2000

A:Accession: J00977; A03143; S42424

R:Hayaaka, K.; Nanao, K.; Tahara, M.; Sato, W.; Takada, G.; Miura, M.; Uyemura, K.

Biochem. Biophys. Res. Commun. 181, 204-207, 1991

A:Title: Isolation and sequence determination of cDNA encoding P2 protein of human perip

A:Reference number: J00977; MUID:92068191

A:Accession: J00977

A:Molecule type: mRNA

A:Residues: 1-132 <HAY>

A:Cross-references: EMBL:X62167; NID:g35185; PIDN:CAA44096.1; PID:g35186

A:Experimental source: fetal spinal cord peripheral myelin

A:Note: authors did not translate the codon for residue 1

R:Suzuki, M.; Kitamura, K.; Sakamoto, Y.; Uyemura, K.

J. Neurochem. 39, 1759-1762, 1982

A:Title: The complete amino acid sequence of human P2 protein.

A:Reference number: A03143; MUID:83058785

A:Accession: A03143

A:Molecule type: protein

A:Residues: 2-98,'N',100-110,'D',112-132 <SUZ>

C:Comment: P2 protein and myelin basic protein together constitute a major fraction of P

C:Genetics:

A:Gene: GDB:PMP2

A:Cross-references: GDB:129030; OMIM:170715

A:Map position: 8q21.3-8q22.1

C:Superfamily: myelin P2 protein

C:Keywords: acetylated amino end; myelin; phosphoprotein

F:2-133/Product: myelin P2 protein #status experimental <MAT>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

F:20/Binding site: phosphate (Tyr) (covalent) #status predicted

F:118-125/Disulfide bonds: #status experimental

Query Match 68.6%; Score 463; DB 1; Length 132;

Best Local Similarity 67.2%; Pred. No. 2.2e-35;

Matches 88; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGVDVIKSESTFKN 60

Db 1 MADAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGVDVIKSESTFKN 60

Db 1 MSNKFLGTWKLVSSENFDDYMKALGVGLATRLKGLNLAQPTVIISKGGDIITIRTESTFKN 60

Qy 61 TEISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120

Db 61 TEISFKLGQEFDEVTADNRKTKSVITLQGSNLQVQRWDGKRETTIKRKLKVGKMWAECKM 120

Qy 121 KGVTSRTRYER 131

Db 121 KGVVCTRIYEK 131

RESULT 8

JH0407

N:myelin P2 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 13-Aug-1999

A:Accession: JH0407

R:Narayanan, V.; Kaestner, K.H.; Tennekoon, G.I.

J. Neurochem. 57, 75-80, 1991

A:Title: Structure of the mouse myelin P2 protein gene.

A:Reference number: JH0407; MUID:91268811

A:Accession: JH0407

A:Molecule type: DNA

A:Residues: 1-132 <NAR>

A:Cross-references: GB:S39508; NID:g232319; PIDN:AAB19249.1; PID:g232320

C:Comment: This basic protein is found in peripheral and central nervous system myel

C:Genetics:

A:Introns: 25/1; 82/3; 116/3

C:Superfamily: myelin P2 protein

Query Match 66.7%; Score 450; DB 2; Length 132;

Best Local Similarity 64.9%; Pred. No. 3.4e-34;

Matches 85; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGVDVIKSESTFKN 60

Db 1 MSNKFLGTWKLVSSENFDDYMKALGVGLNLAQPTVIISKGGDIITIRTESAFKN 60

Qy 61 TEISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120

Db 61 TEISFKLGQEFDEVTADNRKAKSVITLGRSLKQVQKWDGKRETTAIRRLDGRWVEECIM 120

Qy 121 KGVTSRTRYER 131

Db 121 KGVVCTRIYEK 131

RESULT 9

FZHU0

N:fatty acid-binding protein, cardiac and skeletal muscle - human

N:Alternate names: fatty acid-binding protein 3 (FABP3)

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1991 #sequence\_revision 03-May-1996 #text\_change 16-Jul-1999

C:Accession: S15432; JH0199; S00603; I54275; A27248

R:Peeters, R.A.; Veerkamp, J.H.; van Kessel, A.G.; Kanda, T.; Ono, T.

Biochem. J. 276, 203-207, 1991

A:Title: Cloning of the cDNA encoding human skeletal-muscle fatty-acid-binding protei

A:Reference number: S15432; MUID:91248148

A:Accession: S15432

A:Molecule type: mRNA; protein

A:Residues: 1-133 <PEE>

A:Cross-references: EMBL:X56549; NID:g31292; PIDN:CAA39889.1; PID:g31293

A:Experimental source: heart muscle; skeletal muscle

R:Boerchers, T.; Hojrup, P.; Nielsen, S.U.; Roepstorff, P.; Spener, F.; Knudsen, J.

Mol. Cell. Biochem. 98, 127-133, 1990

A:Title: Revision of the amino acid sequence of human heart fatty acid-binding protei

A:Reference number: JH0199; MUID:91094793

A:Accession: JH0199

A:Molecule type: protein

A:Residues: 2-129,'O',131-133 <BOE>

R:Offner, G.D.; Brecher, P.; Sawilovich, W.B.; Costello, C.E.; Troxler, R.F.

Biochem. J. 252, 191-198, 1988



Qy 121 KGVSTRVYER 131  
Db 121 GTAVCTRYEK 131

RESULT 12  
S06479  
fatty acid-binding protein, brain - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 24-Nov-1999  
C:Accession: S06479; JH0198  
R:Schoentgen, F.; Pignede, G.; Bonanno, L.M.; Jolles, P.  
Eur. J. Biochem. 185, 35-40, 1989  
A:Title: Fatty-acid-binding protein from bovine brain. Amino acid sequence and some prop  
A:Reference number: S06479; MUID:90032683  
A:Accession: S06479  
A:Molecule type: protein  
A:Residues: 1-131 <SCH>  
R:Experimental source: brain  
Note: 11-Val, 12-Asp, 14-Lys, 18-Asp, 33-Ala, 34-Ser, 40-Thr, 43-Glu, 44-Lys, 49-Val,  
R:Schoentgen, F.; Bonanno, L.M.; Pignede, G.; Jolles, P.  
Mol. Cell Biochem. 98, 35-39, 1990  
A:Title: Amino acid sequence and some ligand binding properties of fatty acid-binding pr  
A:Reference number: JH0198; MUID:91094811  
A:Contents: annotation  
C:Comment: Synaptosomal Na+-dependent amino acid transport is stimulated by the fatty ac  
C:Superfamily: myelin P2 protein  
C:Keywords: blocked amino end  
F:1/Modified site: blocked amino end (Val) (probably acetylated) #status experimental

Query Match 63.0%; Score 425; DB 2; Length 131;  
Best Local Similarity 62.0%; Pred. No. 6.5e-32;  
Matches 80; Conservative 20; Mismatches 29; Indels 0; Gaps 0;

Qy 3 DAFVGTWKLVSSENFDDYMKVGVGFATPKVAGMAKPNMIISVNGDVITIKSESTFKNE 62  
Db 2 DAFVGTWKLVSSENFDDYMKVGVGFATPKVAGMAKPNMIISVNGDVITIKSESTFKNE 61

Qy 63 ISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 122  
Db 62 ISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 121

Qy 123 VTSTRVYER 131  
Db 122 VVAVRHYEK 130

RESULT 13  
A49184  
fatty acid-binding protein - chicken  
N:Alternate names: lipid-binding protein  
C:Species: Gallus gallus (chicken)  
C:Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999  
C:Accession: A49184; S26599  
R:Godbout, R.  
Exp. Eye Res. 56, 95-106, 1993  
A:Title: Identification and characterization of transcripts present at elevated levels i  
A:Reference number: A49184; MUID:93162137  
A:Accession: A49184  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-132 <GOD>  
A:Cross-references: EMBL:X65459; NID:g63230; PIDN:CAA46451.1; PID:g63231  
A:Experimental source: retina  
A:Note: sequence extracted from NCBI backbone (NCBIP:124757)  
C:Superfamily: myelin P2 protein

Query Match 58.1%; Score 392; DB 2; Length 132;  
Best Local Similarity 56.1%; Pred. No. 6.9e-29;  
Matches 74; Conservative 19; Mismatches 39; Indels 0; Gaps 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATPKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
Db 1 MVEAFCATWKLADSHNFDEYMKALGVGFAMRQGVNVTPTVIISSEGDKVVIRTQSTFKN 60

Qy 61 TEISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120  
Db 61 TEISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120

Qy 121 KGVSTRVYER 132  
Db 121 GDVAVRHYERA 132

RESULT 14  
152524  
testis lipid binding protein - rat  
N:Alternate names: 15 kDa perforatorial protein PERF 15  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 13-Aug-1999  
C:Accession: 152524; I53089  
R:Schmitt, M.C.; Jamison, R.S.; Orgebin-Crist, M.C.; Ong, D.E.  
Biol. Reprod. 51, 239-245, 1994  
A:Title: A novel, testis-specific member of the cellular lipophilic transport protein  
A:Reference number: 152524; MUID:95035569  
A:Accession: 152524  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-132 <RES>  
A:Cross-references: EMBL:U07870; NID:g469051; PIDN:AAA68627.1; PID:g469052  
R:Oko, R.J.; Morales, C.C.; Oko, R.; Morales, C.R.  
Dev. Biol. 166, 235-245, 1994  
A:Title: A novel testicular protein, with sequence similarities to a family of lipid  
A:Reference number: I53089; MUID:95046905  
A:Accession: I53089  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-132 <RE2>  
A:Cross-references: EMBL:U09022; NID:g538268; PIDN:AAA67873.1; PID:g538269  
C:Superfamily: myelin P2 protein

Query Match 57.8%; Score 390; DB 2; Length 132;  
Best Local Similarity 56.5%; Pred. No. 1.1e-28;  
Matches 74; Conservative 26; Mismatches 31; Indels 0; Gaps 0;

Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATPKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
Db 1 MIEPFLGTWKLVSSENFENYRELGVCEPRKVLAKIPKSVISFNGERMDIQAGSACRN 60

Qy 61 TEISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120  
Db 61 TEISFILGQEFDEVTADDRKVKSTITLDGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120

Qy 121 KGVSTRVYER 131  
Db 121 NNWVSTRTYER 131

RESULT 15  
156510  
fatty acid binding protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 13-Aug-1999  
C:Accession: 156510  
R:Bennett, E.; Stenvers, K.L.; Lund, P.K.; Popko, B.  
J. Neurochem. 63, 1616-1624, 1994  
A:Title: Cloning and characterization of a cDNA encoding a novel fatty acid binding p  
A:Reference number: 156510; MUID:95016702  
A:Accession: 156510  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-132 <RES>



A:Cross-references: EMBL:U02096; NID:g476081; PIDN:AAA60455.1; PID:g476082  
C:Superfamily: myelin P2 protein

Query Match	57.0%	Score 385;	DB 2;	Length 132;
Best Local Similarity	55.3%	Pred. No. 3e-28;		
Matches 73;	Conservative 19;	Mismatches 40;	Indels 0;	Gaps 0;

  

QY	1	MCDAFVGTWKLVSSENFDDYMKVGVGFATPKVAGMAKPNMIISVNGDVITIKSESTFKN	60
Db	1	MVDAFCATWKLTDQNFDEYMKALGVGFATQVGNVTKPTVIIISQEGGKVIVRTQCTFKN	60
QY	61	TEISFILGQEFDEVTDADRKVKSTITLDGGVLVHVQKWDGKSTTIKPKREDDKLVVECV	120
Db	61	TEISFQLGEEFEETSIDDRNCKSVIRLDGKLIHVQKWDGKSTTIKPKREDDKLVVECV	120
QY	121	KGVTSTRVYERA	132
Db	121	GDVVAVRCYEKA	132

Search completed: June 7, 2002, 18:20:10  
Job time: 152 sec





Db	1	MADAFVGTWKLVD	SKNFDDYMKSLGVGFAT	QVASMTKPTTIEKNGD	ITITKTQSTFKN	60
QY	61	TEISFILQGFEDV	ADRRKVKSTIILDG	VLVHVQWKGDKSTTKR	KREDDOKLVCEVM	120
Db	61	TEINFQIGFEDV	ADRRKVKSLVTDG	GLKHVQKNGQETLT	RELVDGKLILTH	120
QY	121	KGVTSTRVYER				
Db	121	GSVSTRTYEK				

```

RESULT      2
US-08-470-298B-7
; Sequence 7, Application US/08470298B
; Patent No. 5844081
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: YU, GUO-LIANG
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: CYTOSTATIN I
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850

```

```
Query Match          69.0%; Score 466; DB 2; Length 131;
Best Local Similarity 67.9%; Pred. NO. 6.3e-49;
Matches 89; Conservative 15; Mismatches 27; Indels 0; Gaps 0;
```

Qy	1	MCA	FVGTWKLVSSENFDYMKVEGVGFATRKRVAGMAKNPMIISVNGDVITIKSESTFKN	60
			:           :           :           :           :	
Db	1	MAD	FVGTWKLVDKSNFDDYMKSLGFGFATRQVASMKTPTTIEKNGDITIKTOSTPKN	60
			:           :           :           :           :	
Qy	61	TEI	SFILGQEFDDEVTDADRKKVSTITLDDGLVHVHKWDGKSTTKRKREDDKLVCEVM	120
			:                   :           :       :           :	
Db	61	TEI	NFOLGIEFDDEVTDADRKKVSLVTLDGGKLIHVQKWNGQETTLRELVDCKLILTLH	120
			:           :           :           :           :	
Qy	121	KGV	STRVYER	131
			:	
Db	121	GSV	STRVYEK	131
			:	

```

RESULT      3
US-08-820-825-11
; Sequence 11, Application US/08820825
; Patent No. 5945309
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENZ, REINER L.
; APPLICANT: DILLON, PATRICK
; TITLE OF INVENTION: CYTOSTATIN III
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,825
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF222
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-820-825-11

```

Query Match 69.0%; Score 466; DB 2; Length 133;  
Best Local Similarity 67.9%; Pred. No. 6.4e-49;  
Matches 89; Conservative 15; Mismatches 27; Indels

```

Qy   1  MDAFAVGTWKLVSSENPDDYMKVEGVGFATPKRVAGMAKPNMIISVNGDVITIKSESTFKN 60
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db   1  MADAFVGTWKLVDKSNFDDYMKSLGVGFATRQVASMTKPTTIEKNMGDTITIKTQSTFKN 60
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy   61  TETSFILGQEFDEVTADDRKVKSTFIPLDGGVLHVHGWKGDKSSTIKKRREDDKLVVCEVM 120
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db   61  TEINFLQGIEFDEVTAADRKKVSLVTLDDGGLKHVQKWNGQETTTLRELVDGKLILTLTH 120
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy   121  KGVTSRTRYER 131
       | | | | |
Db   121  GSVVSTRYEK 131

RESULT          4
US-023-073A-7
; Sequence 7, Application US/09023073A
; Patent No. 5977309
; GENERAL INFORMATION:
; APPLICANT: NI, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Rosen, Craig A
; TITLE OF INVENTION: Cytostatin I
; NUMBER OF SEQUENCES: 11
```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,073A
; FILING DATE: 13-FEB-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Wales, Michele M.
; REGISTRATION NUMBER: P-43,975
; REFERENCE/DOCKET NUMBER: PFI75D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5772
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-023-073A-7
;
;
; Query Match 69.0%; Score 466; DB 2;
; Best Local Similarity 67.9%; Pred. No. 6,4e-49;
; Matches 89; Conservative 15; Mismatches 27;
;
QY 1 MDAFVGTWKLVSSENPDYMKVEGVGFATRKVAGMAKPNMIIIS
Db 1 MADAFVGTWKLVDKSNFDDYMKSLGVGFGFATRKVASMTKPTTIEI
QY 61 TEISFILGQEFDEVTADDRKVKSTITLDGGVLVHVQKWGDKSTT
Db 61 TEINFQIGIEFDEVTADDRKVKSLVTLTDGGLKLIHVQKWNGQETT
QY 121 KGVSTSTRVRYER 131
Db 121 GSVVSTRTYEK 131
;
; RESULT 5
; US-09-307-817-11
; Sequence 11, Application US/09307817
; Patent No. 6232291
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER L.
; APPLICANT: DILLON, PATRICK
; TITLE OF INVENTION: CYTOSTATIN III
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/307,817
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/820,825
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: BROOKES, ANDERS A.
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PF222
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 133 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-307-817-11

Query Match          69.0%; Score 466;
Best Local Similarity 67.9%; Pred. No. 6.
Matches      89; Conservative 15; Mismatch

QY  1  MCDAFVGTWKLVSSENFDDYMKYGVGVGFATRKVAGM
      | | | | | | | | | | | | | | | | | | | | | | | |
DB  1  MADAFVGTWKLVDKSNFDDYMKSLGVGVGFATRKVASN
      | | | | | | | | | | | | | | | | | | | | | | | |
QY  61  TEISFTLGGFEEFDEVTADDRKVKSTITLDGGVLHVHQ
      | | | | | | | | | | | | | | | | | | | | | | | |
DB  61  TEINQLGIEFDEVTADDRKVKSLVTLDDGGLIHHV
      | | | | | | | | | | | | | | | | | | | | | | | |
QY  121  KGVTSRTRYER 131
      | | | | | |
DB  121  GSWSTRTRYEK 131
      | | | | | |

RESULT      6
US-09-361-737-7
: Sequence 7, Application US/09361737
: Patent No. 6287812
: GENERAL INFORMATION:
: APPLICANT: Ni, Jian
: APPLICANT: Gentz, Reiner
: APPLICANT: Yu, Guo-Liang
: APPLICANT: Rosen, Craig A
: TITLE OF INVENTION: Cytostatin I
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HUMAN GENOME SCIENCES, INC.
: STREET: 9410 KEY WEST AVENUE
: CITY: ROCKVILLE
: STATE: MARYLAND
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Vers1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/361,737
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/023,073
: FILING DATE: 13-FEB-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Wales, Michele M.

```

REGISTRATION NUMBER: P-43,975  
REFERENCE/DOCKET NUMBER: PF175D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5772  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-361-737-7

Query Match 69.0%; Score 466; DB 4; Length 133;  
Best Local Similarity 67.9%; Pred. No. 6.4e-49;  
Matches 89; Conservative 15; Mismatches 27; Indels 0; Gaps 0;

QY 1 MCDAFVGTWKLVSSENFDDYMKYGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
DB 1 MADAFVGTWKLVSSENFDDYMKYGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
QY 61 TEISFILGOEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120  
DB 61 TEINFLQIEFDEVTADDRKVKSLTLDGGKLIHVQKWDGKSTTIKKRREDDKLVEECVM 120

QY 121 KGVSTSTRVYER 131  
DB 121 GSVVSTRTYER 131

RESULT 7  
US-08-409-731A-11  
Sequence 11, Application US/08409731A  
Patent No. 568758  
GENERAL INFORMATION:  
APPLICANT: Ni, Jian  
APPLICANT: Yu, Guo-Liang  
APPLICANT: Gentz, Reiner  
APPLICANT: Rosen, Craig A.  
TITLE OF INVENTION: CYTOSTATIN I  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
STREET: 9410 KEY WEST AVENUE  
CITY: ROCKVILLE  
STATE: MD  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/409,731A  
FILING DATE: 24-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Robert H  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PF175  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8512  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-409-731A-11

Query Match 68.6%; Score 463; DB 1; Length 132;  
Best Local Similarity 67.2%; Pred. No. 1.5e-48;  
Matches 88; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

QY 1 MCDAFVGTWKLVSSENFDDYMKYGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
DB 1 MSNKFLGTWKLVSSENFDDYMKALGVGLATRKLGKLNLAQPTVIISKGGDIITIRTESTFKN 60  
QY 61 TEISFILGOEFDEVTADDRKVKSTITLDGGVLVHVQKWDGKSTTIKKRREDDKLVEECVM 120  
DB 61 TEISFKLGOEFDEVTADNRKTSIVTLQSGSLNVQVRWDGKETTIIKKRLVNGKVMVAECKM 120

QY 121 KGVSTSTRVYER 131  
DB 121 KGVVCTRIYER 131

RESULT 8  
US-08-470-298B-11  
Sequence 11, Application US/08470298B  
Patent No. 5844081  
GENERAL INFORMATION:  
APPLICANT: NI, JIAN  
APPLICANT: GENTZ, REINER  
APPLICANT: YU, GUO-LIANG  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: CYTOSTATIN I  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
STREET: 9410 KEY WEST AVENUE  
CITY: ROCKVILLE  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,298B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BROOKES, ALLAN A.  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF175D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8512  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: MYELIN P2 (FIGURE 2)  
US-08-470-298B-11

Query Match 68.6%; Score 463; DB 2; Length 132;  
Best Local Similarity 67.2%; Pred. No. 1.5e-48;  
Matches 88; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

QY 1 MCDAFVGTWKLVSSENFDDYMKYGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
DB 1 MSNKFLGTWKLVSSENFDDYMKALGVGLATRKLGKLNLAQPTVIISKGGDIITIRTESTFKN 60

RESULT 10  
US-09-361-737-11

APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: CYTOSTATIN I  
NUMBER OF SEQUENCES: 11

TITLE OF INVENTION: CYTOSTAIN III  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
 STREET: 9410 KEY WEST AVENUE  
 CITY: ROCKVILLE  
 STATE: MD  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/820,825  
 FILING DATE: 19-MAR-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BROOKES, ANDERS A.  
 REGISTRATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: PF222  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504





Db 61 TEISFKLGVEFDETTADDRKVKSIVTLDGGKLVHLQKWGOETTIVRELIDGKLILTLTH 120  
Qy 121 KGVSTSTRVYER 131  
|| ||:  
Db 121 GTAVCTRTYEK 131

Search completed: June 7, 2002, 18:19:47  
Job time: 164 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2002, 18:13:28 ; Search time 30.07 Seconds

(without alignments)  
487.587 Million cell updates/sec

Title: US-09-503-596-4

Perfect score: 675

Sequence: 1 MCDAFVGTWKLVSSENFDDY.....KLVECMVKGVSTRVIERA 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	675	100.0	132	21	AA190320 Human AFABP protei
2	625	92.6	132	21	AA190319 Mouse AFABP protei
3	466	69.0	133	22	AA190318 Mouse MDGI polyep
4	458	67.9	136	19	AA190317 Human myelin P2 pr
5	453	67.1	131	12	AA190316 MDGI active peptid
6	449	66.5	133	22	AA190315 Human MDGI polyep
7	446	66.1	133	22	AA190314 Bovine MDGI polyep
8	445	65.9	136	19	AA190313 Bovine myelin P2 p
9	435.5	64.5	134	18	AA190312 Porcine heart-fatt
10	399	59.1	132	18	AA190311 Human cytoctatin I
11	399	59.1	132	19	AA190310 Amino acid sequenc

12	399	59.1	132	19	AA190309 Human cytoctatin I
13	399	59.1	132	20	AA190308 Human 14-16 kDa FA
14	396	58.7	117	21	AA190307 Human secreted pro
15	386.5	57.3	131	22	AA190306 Human cytoctatin I
16	382	56.6	132	21	AA190305 Human secreted pro
17	381	56.4	132	16	AA190304 Mouse cellular X b
18	375	55.6	88	21	AA190303 Breast and ovarian
19	366	54.2	135	15	AA190302 Melanogenic inhibi
20	366	54.2	135	22	AA190301 Human keratinocyte
21	362	53.6	135	22	AA190300 Mouse keratinocyte
22	361	53.5	158	22	AA190299 Novel human diagno
23	360	53.3	82	22	AA190298 Human colon cancer
24	298.5	44.2	134	22	AA190297 Human NOV10 protei
25	298.5	44.2	163	22	AA190296 Human NOV9 protein
26	297	44.0	132	22	AA190295 Cytoctatin family
27	297	44.0	135	22	AA190294 Cytoctatin family
28	285	42.2	130	22	AA190293 Drosophila melanog
29	270	40.0	69	21	AA190292 Human 5' EST relat
30	237	35.1	134	20	AA190291 Human retinoid bin
31	232.5	34.4	172	22	AA190290 Fatty acid-binding
32	229.5	34.0	133	16	AA190289 Schistosoma mansoni
33	222	32.9	137	14	AA190288 CRABP-I gene produ
34	210	31.1	56	21	AA190287 Human secreted pro
35	201.5	29.9	138	14	AA190286 CRABP-II gene prod
36	201.5	29.9	138	22	AA190285 Human lung tumour
37	201.5	29.9	149	21	AA190284 Lung cancer associ
38	194	28.7	99	21	AA190283 Human secreted pro
39	191.5	28.4	181	22	AA190282 Novel human diagno
40	186	27.6	160	21	AA190281 Human cancer assoc
41	169.5	25.1	176	22	AA190280 Novel human diagno
42	165	24.4	135	22	AA190279 Human polypeptide
43	165	24.4	135	22	AA190278 Human cellular ret
44	164	24.3	168	22	AA190277 Human polypeptide
45	164	24.3	135	18	AA190276 Human cytoctatin I

ALIGNMENTS

RESULT 1  
AA190320  
ID AA190320 standard; Protein; 132 AA.  
XX  
XX  
AC AA190320;  
XX  
DT 22-NOV-2000 (first entry)  
XX  
DE Human AFABP protein sequence.  
XX  
KW AFABP; adipocyte fatty acid binding protein; atherosclerotic lesion;  
KW lesion formation inhibition; macrophage; adipocyte; atherosclerosis;  
KW serum cholesterol; therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2000047734-A1.  
XX  
PD 17-AUG-2000.  
XX  
PF 11-FEB-2000; 2000WO-US03560.  
XX  
PR 12-FEB-1999; 99US-0119880.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Lee M, Perrella MA, Hotamisligil GS;  
XX  
DR WPI: 2000-506094/45.  
DR N-PSDB; AAA37717.  
XX  
PT Reducing expression of adipocyte fatty acid binding protein through  
PT administration of a compound is used to inhibit formation of an  
PT atherosclerotic lesion -

XX Disclosure; Page 14; 43pp; English.  
 XX This sequence represents the human AFABP (adipocyte fatty  
 CC acid binding protein) protein sequence. The invention relates to a method  
 CC for inhibiting formation of an atherosclerotic lesion comprising  
 CC administering to a mammal a compound that reduces expression of adipocyte  
 CC fatty acid binding protein (AFABP). The method is used to inhibit  
 CC formation of atherosclerotic lesions. The method is used to identify  
 CC compounds which can be used to inhibit formation of atherosclerotic  
 CC lesions through inhibition of AFABP binding to an intracellular ligand  
 CC in a macrophage or adipocyte, inhibition of development of an  
 CC atherosclerotic lesion, inhibition of a macrophage differentiating into  
 CC a foam cell or inhibition of AFABP expression in a cell. AFABP activity  
 CC may be inhibited to treat atherosclerosis or to treat individuals at risk  
 CC of developing atherosclerosis. Inhibiting AFABP expression or activity  
 CC reduces the development of atherosclerotic lesions despite a high level  
 CC of serum cholesterol.  
 XX  
 XX Sequence 132 AA;  
 SQ

Query Match 100.0%; Score 675; DB 21; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-71;  
 Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
 Db 1 mcdafvgtwklvsse enfddymkev gvfatrkvagmakpnmiisvngdvi tiksestfkn 60  
 Qy 61 TEISFILGOEFDEVTTADDRKVKSTITLDGVLVHVQKWDGKSTTIKRRKDDKLVVECM 120  
 Db 61 teisfilgqefdevttad drkvvkst itldggvlv hvqkwgskst tikrrkddklvvecm 120  
 Qy 121 KGVTSRVYERA 132  
 Db 121 kgvtstrvyera 132

RESULT 2  
 AAY90319  
 ID AAY90319 standard; Protein; 132 AA.  
 XX  
 AC AAY90319;  
 XX  
 DT 22-NOV-2000 (first entry)  
 XX  
 DE Mouse AFABP protein sequence.  
 XX  
 KW AFABP; adipocyte fatty acid binding protein; atherosclerotic lesion;  
 KW lesion formation inhibition; macrophage; adipocyte; atherosclerosis;  
 KW serum cholesterol; therapy; mouse.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200047734-A1.  
 XX  
 PD 17-AUG-2000.  
 XX  
 PF 11-FEB-2000; 2000WO-US03560.  
 XX  
 PR 12-FEB-1999; 99US-0119880.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 XX Lee M, Ferrella MA, Hotamisligil GS;  
 XX WPI; 2000-506094/45.  
 DR N-PSDB; AAA37716.  
 DR  
 XX Reducing expression of adipocyte fatty acid binding protein through  
 PT administration of a compound is used to inhibit formation of an  
 PT atherosclerotic lesion -

XX Disclosure; Page 14; 43pp; English.  
 XX This sequence represents the mouse AFABP (adipocyte fatty  
 CC acid binding protein) protein sequence. The invention relates to a method  
 CC for inhibiting formation of an atherosclerotic lesion comprising  
 CC administering to a mammal a compound that reduces expression of adipocyte  
 CC fatty acid binding protein (AFABP). The method is used to inhibit  
 CC formation of atherosclerotic lesions. The method is used to identify  
 CC compounds which can be used to inhibit formation of atherosclerotic  
 CC lesions through inhibition of AFABP binding to an intracellular ligand  
 CC in a macrophage or adipocyte, inhibition of development of an  
 CC atherosclerotic lesion, inhibition of a macrophage differentiating into  
 CC a foam cell or inhibition of AFABP expression in a cell. AFABP activity  
 CC may be inhibited to treat atherosclerosis or to treat individuals at risk  
 CC of developing atherosclerosis. Inhibiting AFABP expression or activity  
 CC reduces the development of atherosclerotic lesions despite a high level  
 CC of serum cholesterol.  
 XX  
 XX Sequence 132 AA;  
 SQ

Query Match 92.6%; Score 625; DB 21; Length 132;  
 Best Local Similarity 91.7%; Pred. No. 4.1e-65;  
 Matches 121; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 MCDAFVGTWKLVSSENFDDYMKVGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
 Db 1 mcdafvgtwklvsse enfddymkev gvfatrkvagmakpnmiisvngdvi tiksestfkn 60  
 Qy 61 TEISFILGOEFDEVTTADDRKVKSTITLDGVLVHVQKWDGKSTTIKRRKDDKLVVECM 120  
 Db 61 teisfilgqefdevttad drkvvkst itldggvlv hvqkwgskst tikrrkddklvvecm 120  
 Qy 121 KGVTSRVYERA 132  
 Db 121 kgvtstrvyera 132

RESULT 3  
 AAG66576  
 ID AAG66576 standard; Protein; 133 AA.  
 XX  
 AC AAG66576;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Mouse MDGI polypeptide.  
 XX  
 KW Mouse; cytoostatin III; cytostatic; epithelial cell growth;  
 KW milk production; breast involution; cardiac myocyte hypertrophy;  
 KW leukaemia; MDGI; mammary-derived growth inhibitor.  
 XX  
 OS Mus sp.  
 XX  
 PN US6232291-B1.  
 XX  
 PD 15-MAY-2001.  
 XX  
 PF 10-MAY-1999; 99US-0307817.  
 XX  
 PR 19-MAR-1996; 96US-0013655.  
 PR 19-MAR-1997; 97US-0820825.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX NI J, Yu G, Gentz R, Dillon PJ;  
 XX WPI; 2001-342775/36.  
 XX  
 XX Cytostatin III polypeptides that modulate growth of epithelial cells,  
 PT stimulate milk production in humans and cows and promote involution of  
 PT breast, for research, biological, clinical and therapeutic purposes -





```
Query Match 66.1%; Score 446; DB 22; Length 133;
Best Local Similarity 64.9%; Pred. No. 3.5e-44;
Matches 85; Conservative 15; Mismatches 31; Indels 0; Gaps 0;

QY 1 MCDAFVGTWKLVSSENFDDYMKVEGVGFATRKVAGMAKPMNIISVNGDVITIKSESTPKN 60
DB 1 mvdafvgtwklvsknfdymkslgvfatrgvgnmtkpttiievngdtviiktgstfkn 60

QY 61 TEISFILGOEFDEVTDADRKKVSTITLDGGLVHVHVKWDGKSTTIKKRREDDKLVVECYM 120
DB 61 teisfklgvefdeettadrrkvsivtlldgglvkvhvkwngetslvrenvdgklliltlch 120

QY 121 KGVSTSTRVYER 131
DB 121 gtavctrttyek 131

RESULT 8
AAW40228
AAW40228 standard; Protein; 136 AA.
AAW40228;
26-JUN-1998 (first entry)
DE Bovine myelin P2 protein.
XX Bovine; myelin; P2 protein; treatment; inflammatory polyneuritis;
KW Guillain-Barre syndrome; vasculitis; nerve inflammation;
KW gammopathy.
XX Bos taurus.
OS
XX WO9803647-A2.
PN
XX 29-JAN-1998.
PD
XX 18-JUL-1997; 97WO-DE01535.
PF
XX 18-JUL-1996; 96DE-1029095.
PR
XX (GOLD/) GOLD R.
PA (WEIS/) WEISHAUP A.
XX Gold R, Weishaupt A;
PI
XX WPI; 1998-120772/11.
PT
XX Recombinant myelin proteins for treating T-cell mediated disease of
peripheral nervous system - by high dose antigen therapy, causing
apoptosis in T cells, for treating e.g. polyneuritis or
PT Guillain-Barre syndrome
XX Disclosure; Fig 4; 14pp; German.
XX The present sequence is the bovine myelin P2 protein, which can
CC be used to treat T-cell mediated diseases of the peripheral nervous
CC system, especially chronic-inflammatory polyneuritis,
CC Guillain-Barre syndrome, vasculitis and nerve inflammation in cases
CC of gammopathy.
XX Sequence 136 AA;
SQ

Query Match 65.9%; Score 445; DB 19; Length 136;
Best Local Similarity 66.2%; Pred. No. 4.7e-44;
Matches 86; Conservative 17; Mismatches 27; Indels 0; Gaps 0;

QY 1 MCDAFVGTWKLVSSENFDDYMKVEGVGFATRKVAGMAKPMNIISVNGDVITIKSESTPKN 60
DB 1 mskkfkgtwklvssenfddymkalgvlatrklglnlakppviiskkgdiitrtspfn 60
```

```
QY 61 TEISFILGOEFDEVTDADRKKVSTITLDGGLVHVHVKWDGKSTTIKKRREDDKLVVECYM 120
DB 61 teisfklgvefdeettadrrkvsivtlldgglvkvhvkwngetslvrenvdgklliltlch 120

QY 121 KGVSTSTRVYER 130
DB 121 kvvctrttyd 130

RESULT 9
AAW31534
AAW31534 standard; Protein; 134 AA.
XX
AC AAW31534;
XX
DT 20-APR-1998 (first entry)
XX Porcine heart-fatty acid binding protein.
DE Porcine; pig; swine; heart-fatty acid binding protein; H-FABP;
KW localisation; identification; marking; gene; allele;
KW quantitative trait locus; production trait; transgenic animal;
KW fat reduction.
XX
OS Sus scrofa.
XX WO9735878-A2.
PN
XX 02-OCT-1997.
PD
XX 27-MAR-1997; 97WO-NL00157.
PF
XX 28-MAR-1996; 96EP-0200855.
PR
XX (DALL-) DALLAND BV.
PA (DIER-) INST DIERHOUDERIJ EN DIERGEZONDHEID ID-D.
PA (NNEV-) NOORD NEDERLANDS VARKENSTAMBOEK BV.
PA (PROV-) PROVA BV.
PA (STAN-) STAMBOEK ZUID BV.
XX
XX Gerbens F;
XX WPI; 1997-489563/45.
XX N-PSDB; AAT89493.
XX Pig heart fatty acid-binding protein gene - used to identify
PT polymorphisms associated with production traits, e.g. body weight,
PT intramuscular fat
XX Claim 1; Fig 1; 41pp; English.
PS
XX The present sequence is porcine heart-fatty acid binding
CC protein (H-FABP). The H-FABP gene can be used to localise, identify
CC or mark porcine genes, alleles or quantitative trait loci,
CC especially where these are associated with production traits.
CC Alleles of the porcine H-FABP gene can be marked, allowing them to
CC be distinguished, preferably by detection of specific restriction
CC sites, e.g. MspI, HaeIII or HinfI. The H-FABP gene can also be used
CC to localise, identify or mark genes, alleles or quantitative trait
CC loci in samples by amplification of specific genomic fragments. By
CC using samples large numbers of pigs can be genotyped rapidly. The
CC methods can be used to identify differences between pig alleles
CC associated with improvements in production traits, and in marker
CC assisted identification/selection of pigs. The H-FABP gene can also
CC be used to generate transgenic animals encoding desirable alleles
CC of H-FABP, to enhance production or physiological characteristics.
CC Specific allelic proteins or peptides may also be produced from the
CC gene. Such peptides, or antibodies directed against them, can be
CC used to influence production traits in live animals or cell
CC culture, or in diagnostic tests to select animals expressing
CC desired forms of allelic proteins/fragments. Fat reduction is
CC desirable in pig breeding because of interest in lean meat, but
CC previous efforts to reduce backfat deposits have also decreased
```





CC adenocarcinoma, in particular, cancers of neurological origin such as  
CC cancers of the breast and kidney. The products can also be used for  
CC detection, diagnosis and drug screening.  
XX  
SQ Sequence 132 AA;  
  
Query Match 59.1%; Score 399; DB 19; Length 132;  
Best Local Similarity 56.8%; Pred. No. 1.1e-38;  
Matches 75; Conservative 21; Mismatches 36; Indels 0; Gaps 0;  
  
Qy 1 MCDAFVGTWKLVSSENFDDYKKEVGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
Db 1 mveafcatwkltnsqnfdeymkalgvftrqvgntvktptviisqegdkvvtlrfstfkn 60  
  
Qy 61 TEISFILGOEFDEVTDADRKYKSTITLDGGLVHVQKWDGKSTTIKRREDDKLVVECYM 120  
Db 61 teisfqlgeefdettdadrncksvvslgdgklvhiqkwdgktnfvreikdgkmvmtltf 120  
  
Qy 121 KGVTSRTRYERA 132  
Db 121 gdvavvrhyeka 132  
  
RESULT 12  
AAW81106  
ID AAW81106 standard; Protein; 132 AA.  
AC AAW81106;  
DT 29-JAN-1999 (first entry)  
DE Human cytotstatin II.  
KW Human; cytotstatin; cell growth; tumour; nervous system;  
KW viral infection; microbial infection.  
XX Homo sapiens.  
XX WO9844109-A1.  
PD 08-OCT-1998.  
PE 25-MAR-1998; 98WO-US05839.  
PR 27-MAR-1997; 97US-0041645.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (LONG-) LONG ISLAND JEWISH MEDICAL CENT.  
Gentz RL, Nardelli B, Ni J, Shi YE, Yu G;  
WPI; 1998-557110/47.  
DR N-PSDB; AAV68225.  
PT New isolated human cytotstatin II - used to develop products for the  
PT treatment of e.g. cancers or viral or microbial infections or for  
PT protecting nervous system cells from toxic agents  
XX  
PS Claim 1; Fig 1; 73pp; English.  
XX  
CC The nucleotide sequence encoding Human cytotstatin can inhibit cell growth  
CC and modulate differentiation. The cytotstatin II polypeptides can be used  
CC for inhibiting tumour growth in a subject, for stimulating growth of or  
CC protecting nervous system cells from toxic agents or for protecting  
CC against or treating viral or microbial infections in mammals. The  
CC products can also be used e.g. to modulate angiogenesis, to modulate  
CC breast development and milk production. They can also be used in  
CC cerebella granular cells and photo receptor cells to provide protection  
CC from lipid peroxidation associated with the oxidative stress induced  
CC during early stages of ischemia, apoptosis, and excitatory amino acid  
CC induced cell death. The retinoid binding potential of cytotstatin II  
CC may be used on photo receptor cells in vivo or in vitro. The activity of

CC haematopoiesis indicates a possible immunosuppressive activity or a  
CC lineage specific stimulation of haematopoiesis which could be used for  
CC treating conditions requiring immunosuppression. Antagonists to  
CC cytotstatin II may be used in vivo to induce deficiencies or enhancement  
CC in the immune or in the haematopoietic systems. They may be used e.g. to  
CC treat cardiac myocyte hypertrophy or leukemia.  
XX  
SQ Sequence 132 AA;  
  
Query Match 59.1%; Score 399; DB 19; Length 132;  
Best Local Similarity 56.8%; Pred. No. 1.1e-38;  
Matches 75; Conservative 21; Mismatches 36; Indels 0; Gaps 0;  
  
Qy 1 MCDAFVGTWKLVSSENFDDYKKEVGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
Db 1 mveafcatwkltnsqnfdeymkalgvftrqvgntvktptviisqegdkvvtlrfstfkn 60  
  
Qy 61 TEISFILGOEFDEVTDADRKYKSTITLDGGLVHVQKWDGKSTTIKRREDDKLVVECYM 120  
Db 61 teisfqlgeefdettdadrncksvvslgdgklvhiqkwdgktnfvreikdgkmvmtltf 120  
  
Qy 121 KGVTSRTRYERA 132  
Db 121 gdvavvrhyeka 132  
  
RESULT 13  
AAW82403  
ID AAW82403 standard; Protein; 132 AA.  
AC AAW82403;  
DT 23-FEB-1999 (first entry)  
DE Human 14-16 kDa FABP analogue GEN 128B10 protein.  
KW FABP; detection; gene expression; analogue; GEN 128B10.  
XX Homo sapiens.  
XX JP10286089-A.  
PD 27-OCT-1998.  
PE 15-APR-1997; 97JP-0096908.  
PR 15-APR-1997; 97JP-0096908.  
XX (SAKA ) OTSUKA PHARM CO LTD.  
XX WPI; 1999-063481/06.  
DR N-PSDB; AAV73482, AAV73483.  
XX  
PT New human rab7GTP-combined analogous protein gene - useful for  
PT detection of its expression in tissues  
XX  
PS Example 3; Page 24; 35pp; Japanese.  
XX  
CC This sequence represents a novel human 14-16 kDa FABP protein analogue,  
CC GEN 128B10. The gene is useful for the detection of gene expression in  
CC various tissues.  
XX  
SQ Sequence 132 AA;  
  
Query Match 59.1%; Score 399; DB 20; Length 132;  
Best Local Similarity 56.8%; Pred. No. 1.1e-38;  
Matches 75; Conservative 21; Mismatches 36; Indels 0; Gaps 0;  
  
Qy 1 MCDAFVGTWKLVSSENFDDYKKEVGVGFATRKVAGMAKPNMIISVNGDVITIKSESTFKN 60  
Db 1 mveafcatwkltnsqnfdeymkalgvftrqvgntvktptviisqegdkvvtlrfstfkn 60



Job time: 360 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:32:01 ; Search time 2892.64 Seconds  
(without alignments)  
2416.970 Million cell updates/sec

Title: us-09-503-596-8  
Perfect score: 518  
Sequence: 1 gaattccacaggaatcagg.....ccctgggtcttctctcttaga 518

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.6	10.7	997	12 CNS005TE	AL060767 Drosophil
2	54.2	10.5	1101	12 CNS00F7H	AL070235 Drosophil
3	52.8	10.2	1044	12 AG056564	AG056564 Pan trogl
4	47.6	9.2	987	12 CNS00418	AL066537 Drosophil
5	47	9.1	774	12 CNS01FGX	AL141826 Anopheles
6	45.6	8.8	1101	12 CNS017RP	AL108415 Drosophil
7	45.6	8.8	1123	10 BG701181	BG701181 602680785
8	45.4	8.8	1031	9 BE036673	BE036673 MP06E09 M
9	44.6	8.6	1246	12 AG135261	AG135261 Pan trogl
10	44.4	8.6	667	12 AZ370954	AZ370954 LM0122K16
11	44.4	8.6	959	12 CNS008MZ	AL052079 Drosophil
12	44.2	8.5	875	10 BF264737	BF264737 HV_CEA001
13	44.2	8.5	991	10 BM414939	BM414939 OP20007 M
14	43.8	8.5	621	10 BI390101	BI390101 ppp1c.pk0
15	43.8	8.5	1322	10 BE966081	BE966081 601659991
16	43.6	8.4	1101	12 CNS0181N	AL108773 Drosophil
17	43.4	8.4	322	12 CNS00HDD	AL073359 Drosophil

C 18	43.4	8.4	412	12 AQ557203	AQ557203 HS_2081_A
C 19	43.4	8.4	877	12 CNS00D37	AL060573 Drosophil
C 20	43.4	8.4	1101	12 CNS002BV	AL097477 Drosophil
C 21	43.2	8.3	1080	12 AG130798	AG130798 Pan trogl
C 22	43	8.3	539	9 AV685047	AV685047 AV685047
C 23	43	8.3	561	12 AZ886506	AZ886506 RPCI-23-1
C 24	42.8	8.3	389	12 CNS03YLI	AL266463 Tetraodon
C 25	42.8	8.3	477	12 AQ390696	AQ390696 CITHI-El-
C 26	42.8	8.3	586	12 AQ390699	AQ390699 CITHI-El-
C 27	42.8	8.3	1152	12 AQ779752	AQ779752 HS_5572_B
C 28	42.6	8.2	590	12 BG112597	BH012597 TDGAL50TH
C 29	42.6	8.2	787	12 BH613790	BG613790 602639777
C 30	42.4	8.2	407	12 AZ083954	AZ083954 RPCI-23-4
C 31	42.4	8.2	580	10 BE274509	BE274509 601120367
C 32	42.4	8.2	715	10 BG495803	AG495803 602539064
C 33	42.4	8.2	1000	12 AG042808	AG042808 Pan trogl
C 34	42.2	8.1	497	9 AI222742	AI222742 qp38d10.x
C 35	42.2	8.1	565	10 BM439556	BM439556 pgr1c.pk0
C 36	42.2	8.1	592	10 BI392756	BI392756 pgr1c.pk0
C 37	42.2	8.1	611	9 AL588140	AL588140 AL588140
C 38	42	8.1	543	12 AQ500504	AQ500504 V38A2 mTn
C 39	42	8.1	597	9 AV762619	AV762619 AV762619
C 40	42	8.1	743	12 B20664	B20664 T408-T7 TAM
C 41	42	8.1	971	12 CNS006JC	AL065539 Drosophil
C 42	42	8.1	1101	12 CNS007MD	AL067247 Drosophil
C 43	41.8	8.1	520	12 AZ365733	AZ365733 IM0112H10
C 44	41.8	8.1	1101	12 CNS00F86	AL070260 Drosophil
C 45	41.6	8.0	923	12 AZ205171	AZ205171 SP_0100_A

## ALIGNMENTS

### RESULT 1

LOCUS CNS005TE 997 bp DNA linear GSS 03-JUN-1999  
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #  
BAC12K22 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL060767  
VERSION AL060767.1 GI:4943573

KEYWORDS GSS  
SOURCE fruit fly.

### ORGANISM

Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 997)

### REFERENCE

### AUTHORS

### TITLE

### JOURNAL

### COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mamoser in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
ECORI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
p1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

### FEATURES

source

1. 997  
/organism="Drosophila melanogaster"  
/db\_xref="taxon:7227"  
/clone\_lib="RPCI-98"











Seq primer: CGTTGTAACAGCGCCAGT  
Class: plasmid ends

High quality sequence stop: 667.

Location/Qualifiers

## FEATURES

source

1. .667

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="JUGC1M0122K16"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="F. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g1/473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

134 a 177 c 101 g 254 t 1 others

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 108; Conservative

8.6%; Score 44.4; DB 12; Length 667;

Mismatches 0; Indels 0; Gaps 0;

QY 300 aggggtggagccactgtggtgagcactctctatggctccctttctgtatttctat 359

Db 414 AGGGGAGGCACGCTGCTTACATTCAGTGAGTTCTTCTTCTTCTTCTTCTTCT 473

QY 360 gttttctgagctcttttcccgcttattgatttcttttttttttttttttttttt 419

Db 474 TCTTCT 533

QY 420 cctctctctatataatgctcctcaggggttcattctgaatcatctactgtgaactattc 479

Db 534 NCT 593

QY 480 ccattgtttgcagagcccccctggtctctcttc 514

Db 594 TTATTATTATATGTAAGTACACTGAGTGCTCTC 628

RESULT 11

CNS008MZ

LOCUS

DEFINITION

959 bp DNA linear GSS 03-JUN-1999

Drosophila melanogaster genome survey sequence TET3 end of BAC #

BAC17K04 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

ACCESSION

AL052079

VERSION

AL052079.1

KEYWORDS

GSS.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 959)

Genoscope.

Direct Submission

TITLE

Submission

## JOURNAL

## COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osoegawa and  
Aaron Mammos in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

source

1. .959

/organism="Drosophila melanogaster"

/db\_xref="taxon:7227"

/clone\_lib="RPCI-98"

/clone="BAC17K04"

/note="end : TET3"

BASE COUNT 178 a 255 c 138 g 261 t 127 others

ORIGIN

## Query Match

Best Local Similarity

Matches 85; Conservative

8.6%; Score 44.4; DB 12; Length 959;

Mismatches 66; Indels 0; Gaps 0;

QY 195 qtggaagtgtcacagcccaaacctcccaagctcagccctcttgccttgtaacaa 254

Db 656 GGGSGKCC 715

QY 255 tcaagcgcctctggaactgcgcgcctctctcttcttggcagggttgagccac 314

Db 716 YCCYTCCTCCMCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 775

QY 315 tttgctgagcagctcttctgctcttcttcttcttcttcttcttcttcttctt 374

Db 776 CCYCYCYHY 835

QY 375 ttcgccgccttctgattcttcttcttcttcttcttcttcttcttcttctt 434

Db 836 TCYTCYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTY 895

QY 435 tgcctctcaggttcttcttcttcttcttcttcttcttcttcttcttcttctt 481

Db 896 TTTYVVCYCYV 942

## RESULT 12

BF264737/c

LOCUS

DEFINITION

875 bp mRNA linear EST 23-OCT-2001

HV\_CEA0010D19f Hordeum vulgare seedling green leaf EST library

HV\_CEA0010D19f, mRNA sequence.

ACCESSION

BF264737

VERSION

BF264737.1

KEYWORDS

EST.

SOURCE

barley.

ORGANISM

Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

; Triticeae; Hordeum.

REFERENCE

1 (bases 1 to 875)

Wing, R., Close, T. J., Kleinhofs, A., Wise, R., Wei, F., Begum, D.,

Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi

, D.W., Fenton, R.D., Oates, R. and Main, D.

Development of a genetically and physically anchored EST resource





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:33:56 ; Search time 102.09 Seconds  
(without alignments)  
1246.333 Million cell updates/sec

Title: US-09-503-596-8  
Perfect score: 518  
Sequence: 1 gaattccagcaggaatcagg.....ccctgggtcttctcttaga 518

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents.NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	518	1	US-07-771-022F-1
2	508.8	98.2	512	5	PCT-US91-06989-1
3	74.2	14.3	7218	1	US-08-232-463-14
4	41	7.9	41	1	US-07-771-022F-16
5	40.6	7.8	72604	4	US-09-268-992-7
6	39.6	7.6	6042	1	US-08-261-822A-1
7	39.6	7.6	6042	5	PCT-US95-07744A-1
8	39.6	7.6	6172	2	US-08-819-288-1
9	39.6	7.6	6172	4	US-09-400-348-1
10	38.8	7.5	289	4	US-09-007-005-17
11	38.8	7.5	289	4	US-09-244-796-17
12	38.2	7.4	372	4	US-09-018-584A-13
13	38.2	7.4	152331	3	US-09-128-155-16
14	38.2	7.4	176373	3	US-09-128-155-17
15	36.6	7.1	5433	3	US-08-929-329-1
16	36.4	7.0	2389	3	US-08-691-563C-52
17	36	6.9	333	4	US-09-018-584A-27
18	35.6	6.9	12047	2	US-09-022-461-1
19	35.6	6.9	29604	3	US-08-781-891-207
20	35.4	6.8	597	2	US-08-332-766A-19
21	35	6.8	6755	3	US-08-931-995-4
22	34.8	6.7	454	2	US-08-623-908A-6
23	34.6	6.7	291	1	US-07-922-723A-7
24	34.6	6.7	291	1	US-07-799-828C-7
25	34.6	6.7	291	1	US-08-074-275-7
26	34.6	6.7	291	1	US-08-480-366-7
27	34.6	6.7	291	2	US-07-952-277A-7

C 28	34.4	6.6	415	4	US-09-018-584A-18	Sequence 18, Appl
C 29	34.2	6.6	243	1	US-07-922-723A-9	Sequence 9, Appl
C 30	34.2	6.6	243	1	US-07-799-828C-9	Sequence 9, Appl
C 31	34.2	6.6	243	1	US-08-074-275-9	Sequence 9, Appl
C 32	34.2	6.6	243	1	US-08-480-366-9	Sequence 9, Appl
C 33	34.2	6.6	243	2	US-07-952-277A-9	Sequence 9, Appl
C 34	33.4	6.4	3892	2	US-08-555-723B-3	Sequence 3, Appl
C 35	33.4	6.4	3892	3	US-09-123-465-3	Sequence 3, Appl
C 36	33	6.4	33	1	US-07-771-022F-15	Sequence 15, Appl
C 37	32.8	6.3	1859	3	US-08-691-563C-46	Sequence 46, Appl
C 38	32.4	6.3	569	2	US-08-485-778-5	Sequence 5, Appl
C 39	32.4	6.3	2791	4	US-09-570-367C-1	Sequence 1, Appl
C 40	32.4	6.3	80246	4	US-09-078-294-4	Sequence 4, Appl
C 41	32.4	6.3	80595	4	US-09-078-294-3	Sequence 3, Appl
C 42	32.2	6.2	394	2	US-08-623-908A-7	Sequence 7, Appl
C 43	32.2	6.2	533	6	5482709-5	Patent No. 5482709
C 44	32.2	6.2	152331	3	US-09-128-155-16	Sequence 16, Appl
C 45	32	6.2	1559	4	US-09-019-095A-7	Sequence 7, Appl

## ALIGNMENTS

## RESULT 1

US-07-771-022F-1

; Sequence 1, Application US/07771022F

; Patent No. 5476926

; GENERAL INFORMATION:

; APPLICANT: SPIEGELMAN, BRUCE M.

; APPLICANT: GRAVES, REED A.

; APPLICANT: ROSS, SUSAN R.

; TITLE OF INVENTION: Adipocyte-Specific DNA

; TITLE OF INVENTION: Sequences And Use Thereof In The Production Of

; TITLE OF INVENTION: Transgenic Animals Exhibiting Altered Fat

; TITLE OF INVENTION: Tissue Metabolism.

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: The Dana-Farber Cancer Institute

; STREET: 44 Binney Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02115

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1 MB storage

; COMPUTER: IBM-compatible Compaq Prolinea MT 4/66

; OPERATING SYSTEM: MS/DOS

; SOFTWARE: WordPerfect 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07771,022F

; FILING DATE: 07-OCT-1991

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/589,971

; FILING DATE: 28-SEP-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: HART, JULIA D.

; REGISTRATION NUMBER: 33132

; REFERENCE/DOCKET NUMBER: DFCI-117

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 632-4016

; TELEFAX: (617) 632-4012

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 518 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; HYPOTHETICAL: no

; ANTI-SENSE: no

; ORIGINAL SOURCE:

; ORGANISM: Mus musculus

```
; STRAIN: Swiss Webster
; DEVELOPMENTAL STAGE: embryonic
; TISSUE TYPE: embryonic fibroblast
; CELL TYPE: fibroblast
; CELL LINE: 3T3-F442A
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: ap2911
; US-07-771-022F-1

Query Match      100.0%; Score 518; DB 1; Length 518;
Best Local Similarity 100.0%; Pred No. 1.4e-151;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattccagcaggaatcaggtagctggagaatcgacagagccatcgattcttggcaag 60
Db 1 GAATTCAGCAGGAATCAGGTAGCTGGAGATCGCAGAGCCATCGGATCTTGGCAAG 60

QY 61 ccatgcacaaaggcagaatgacatttccaccagagagaaggattgatgtcagcag 120
Db 61 CCATGCACAAAGGCAGAAATGACATTTCCACCAGAGAGAGGATGATGTGTCAGCAG 120

QY 121 aagtcaccacccagagacaaatggagttccagatgcctgacatttgcttcttactgg 180
Db 121 AAGTCACCACCCAGAGCAAAATGGAGTTCCACAGATGCCTGACATTTGCCTTCTTACTGG 180

QY 181 atcagagttcactagtggaagtgtcacagcccaaacactcccccaagctcagcccttc 240
Db 181 ATCAGAGTTTCACTAGTGAAGTGTACAGCCCAAAACACTCCCCCAAGCTCAGCCCTTCC 240

QY 241 tgcctctgaacaataaaccgctcctgtagtaactgtccgcccctctgtctcttggca 300
Db 241 TGCCCTTGTAACAATAACGCGCTCTGTGATGAACCTGCTCCGCCCTCTCTCTTTGGCA 300

QY 301 gggttgagccactgtgacctgtatgtctatgtctcttctgtctcttctgtgatttcag 360
Db 301 GGTTTGGAGCCCACTGTGCGCTGAGGACTTCTATGGCTCCCTTTTCTGTGATTTTCATG 360

QY 361 gttctgagctcttctcccccgtttatgatttctcttcttcttctctctcttgetaaac 420
Db 361 GTTCTGAGCTCTTTTCCCGCTTTATGATTTTCTCTTTTGTCTCTCTCTCTGCTAAAC 420

QY 421 ctctctgtatatagtccctctcaggttcttcttctgctctcttctcttctcttctct 480
Db 421 CTCCTCTGATATATGCCCCTCAGGTTTCATTTCTGAAATCATCTACTGTGAACATATCC 480

QY 481 cattgttgccagaagcccccgttggttcttctctctctctctctctctctctctct 518
Db 481 CATTGTTTCCAGAGCCCGCTGTTCTTCTCTCTTAGA 518

RESULT 2
PCT-US91-06989-1
; Sequence 1, Application PC/TUS9106989
; GENERAL INFORMATION:
; APPLICANT: SPIEGELMAN, BRUCE M.
; APPLICANT: GRAVES, REED A.
; APPLICANT: ROSS, SUSAN R.
; TITLE OF INVENTION: Adipocyte-Specific DNA
; TITLE OF INVENTION: Sequences And Use Thereof In The Production Of
; TITLE OF INVENTION: Transgenic Animals Exhibiting Altered Fat Metabolism.
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Dana-Farber Cancer Institute
; STREET: 44 Binney Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02115
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720Kb
; MEDIUM TYPE: storage
```

```
; COMPUTER: IBM Personal System 2; Model 30
; OPERATING SYSTEM: MS/DOS
; SOFTWARE: WordPerfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/06989
; FILING DATE: 19910925
; CLASSIFICATION: 380
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/589,971
; FILING DATE: 28-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: HART, JULIA D.
; REGISTRATION NUMBER: 33132
; REFERENCE/DOCKET NUMBER: DFCI-117 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255-8900
; TELEFAX: (203) 259-2846
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; STRAIN: Swiss Webster
; DEVELOPMENTAL STAGE: embryonic
; TISSUE TYPE: embryonic fibroblast
; CELL TYPE: fibroblast
; CELL LINE: 3T3-F442A
; IMMEDIATE SOURCE:
; LIBRARY: genomic
; CLONE: ap2911
; PCT-US91-06989-1
```

```
Query Match      98.2%; Score 508.8; DB 5; Length 512;
Best Local Similarity 98.6%; Pred No. 9.8e-149;
Matches 510; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 ccagcagggaatcaggtagctggagaatcgacagagccatcgattcttggcaagcag 65
Db 1 CCAGCAGGAATCAGGTAGCTGGAGAATCGCAGAGCCATCGGATCTTGGCAAGCAG 60

QY 66 cgaacaaaggcagaatgcacatttcaccacagagagaaggattgatgtcagcaggaagtc 125
Db 61 CGACAAAGGCAGAAATGCACATTTCCACCAGAGAGAGGATTTAGTCAGCAGGAAGTC 120

QY 126 accaccagagagcaaatggagttccagatgcctgacattgaccttcttactgactcag 185
Db 121 ACCACCCAGAGAGCAAAATGGAGTTCCAGATGCCCTGACATTTGCCCTTCTTACTGGATCAG 180

QY 186 agttcactagtgaagtgtcacagcccaaacactcccccaagctcagcccttcttgc 245
Db 181 AGTTCACTAGTGAAGTGTACAGGCCCAAAACACTCCCCCAAGCTCAGCCCTTCTTGGCC 240

QY 246 ttgtaacaatacagccgctccttgatgaactgtcgcctctctgtctcttcttggcagggtt 305
Db 241 TTGTAACAATCAAGCCGCTCTCTGGATGAACCTCGCCCTCTCTCTCTCTTTTGGCAGGGTT 300

QY 306 ggagccactgtgacctgagcgaactctatggctcccttctctgtgatttctcaggttctc 365
Db 301 GGAGCCCACTGTGGCTGAGCGACTTCTATGGCTCCCTTTTCTGTGATTTTTCATGGTTTC 360

QY 366 tgagctctttcccccgctttatgatatttctcttcttctctctctctctctctctctctct 425
Db 361 TGAGCTCTTTTCCCCGCTTTATGATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

QY 426 tcgtatatgcctctcaggtttcatttctgaatcatctactgtgaactatttccattg 485
Db 426 TCGTATATGCCTCTCAGGTTTCAATTTCTGAATCATCTACTGTGAACATATTCCTCTCTCT 485
```

Qy	323	gagcactctatgcgtcccttcttgtagtttcattcatggttctcgacctttccccccc	382
:	:	: : : : :	:
Db	1224	YVYY	1283
Qy	383	cttatgatcttctcttttgtctctctctgaacacctctctgtatatataaccctct	442
:	:	: : : : :	:
Db	1284	YVYY	1343
Qy	443	caggttctattcagtcatctactgtgaactatgccattctccattgttgccagaagccccc	502
:	:	: : : : :	:
Db	1344	YVYY	1403
Qy	503	g9ttctctctet 515	
:	:	: : : : ::	:
Db	1404	YVYYYYYYYYYYY 1416	
<b>RESULT     4</b>			
US-07-771-022F-16			
; Sequence 16, Application US/0771022F			
; Patent No. 5476926			
; GENERAL INFORMATION:			
; APPLICANT: SPIEGELMAN, BRUCE M.			
; APPLICANT: GRAVES, REED A.			
; APPLICANT: ROSS, SUSAN R.			
; TITLE OF INVENTION: Adipocyte-Specific DNA			
; TITLE OF INVENTION: Sequences And Use Thereof In The Production Of			
; TITLE OF INVENTION: Transgenic Animals Exhibiting Altered Fat			
; TITLE OF INVENTION: Tissue Metabolism.			
; NUMBER OF SEQUENCES: 16			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: The Dana-Farber Cancer Institute			
; STREET: 44 Binney Street			
; CITY: Boston			
; STATE: Massachusetts			
; COUNTRY: USA			
; ZIP: 02115			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette, 3.50 inch, 1 MB storage			
COMPUTER: IBM-compatible Compaq Prolinea MT 4/66			
OPERATING SYSTEM: MS/DOS			
SOFTWARE: WordPerfect 6.0			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/07/771,022F			
FILING DATE: 07-OCT-1991			
CLASSIFICATION: 800			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 07/589,971			
FILING DATE: 28-SEP-1990			
ATTORNEY/AGENT INFORMATION:			
NAME: HART, JULIA D.			
REFERENCE/DOCKET NUMBER: DFCEI-117			
REGISTRATION NUMBER: 33132			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (617) 632-4016			
TELEFAX: (617) 632-4012			
INFORMATION FOR SEQ ID NO: 16:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 41 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: genomic DNA			
HYPOTHETICAL: no			
ANTI-SENSE:			
ORIGINAL SOURCE:			
ORGANISM: Mus musculus			
STRAIN: Swiss Webster			
DEVELOPMENTAL STAGE: embryonic			
TISSUE TYPE: embryonic fibroblast			
CELL TYPE: fibroblast			
CELL LINE: 3T3-F442A			
IMMEDIATE SOURCE:			

```

: LIBRARY: genomic
: CLONE: ap2911
: FEATURE:
: NAME/KEY: ARE7-containing nucleotide sequence
: LOCATION: 29 - 41
: OTHER INFORMATION: Nucleotides 29 to 41 of SEQ ID NO 16
: OTHER INFORMATION: are the nucleotides on
: OTHER INFORMATION: the coding strand of the mouse ap2 gene that are complementary
: OTHER INFORMATION: nucleotides 9 - 19 of SEQ ID
: OTHER INFORMATION: NO 9, which are on the non-coding strand of the mouse ap2 gene
: PUBLICATION INFORMATION:
: AUTHORS: GRAVES, REED A.
: AUTHORS: TONTOZ, PETER
: AUTHORS: SPEIGELMAN, BRUCE M.
: TITLE: Analysis of a Tissue-Specific Enhancer: ARF6
: TITLE: Regulated Adipogenic Gene Expression
: JOURNAL: Molecular and Cellular Biology
: VOLUME: 12
: PAGES: 1202-1208
: DATE: 00-MAR-1992
: IS-07-771-022F-16

```

Query Match	7.9%	Score 41;	DB 1;	Length 41;
Best Local Similarity	100.0%	Pred. No. 0.00055;		
Matches 41;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	139	caaatggaggtccagatgacctgacatttgcctcttactg	179	
db	1	CAAAATGGAGTTCACAGATGCCTGACATTCGCTTCTTACTG	41	

```

RESULT      5
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Feilner, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

```

Query Match	7.8%	Score 40.6;	DB 4;	Length 72604;
Best Local Similarity	52.0%;	Pred. No. 0.026;		
Matches	91; Conservative	0; Mismatches	84; Indels	0; Gaps
339	ttcccttttcgtgattttcatggtcttcagagctcttttcccccggatttatgatttctctct			398
45334	TCCTTTCCTCTTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT			45275

[illegible]

```

RESULT 6
US-08-261-822A-1
; Sequence 1, Application US/08261822A
; Patent No. 5650553
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph R. et al.
; TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
; TITLE OF INVENTION: and Pathogens
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,822A
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
US-08-261-822A-1

```

Query Match	7.6%	Score 39.6	DB 1	Length 6042
Best Local Similarity	56.0%	Pred. No. 0.016		
Matches 75	Conservative 0	Mismatches 59	Indels 0	Gaps 0
339	tcacattttctgtgatttttcattggtttcttgagctcttttccccgcgctttatgattttctct	398		
118	TCCTTTATCTCTCTCTTTTCATGGAACTGAGCTTCTTCTCTTTCTCTCTCTCTCTCT	177		
399	ttttgtctctctctgttgtaaacctctctcgatatatatgcctctcagggtttcattctga	458		
178	CCTCATCTCTATCTCTCGTAGCTTGATAGAGGTTTCTCTCTTTTGAGATCCGTTTCTCT	237		
459	atcatctactgtga	472		
238	CTCTCTCACTGAGA	251		

RESULT 7  
PPCT-US95-07744A-1  
; Sequence 1, Application PC/TUS9507744A  
; GENERAL INFORMATION:



```

Query Match          7.6%; Score 39.6; DB 5; Length 6042;
Best Local Similarity 56.0%; Pred. No. 0.016;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 339 tcccttttctgtatttcacgtgttcttgagctctttttccccccgctttatgattttctct 398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 TCCTTATCTCTCTCTCTTCGATGGAAGCTGCTCTTCTCTCTTTCGCTCTTCTCTTCTCT 177
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 ttgttctctctctctgtaaacctccctctgatatatgcctctcaggtttcattctga 458
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 CTCTATCTCTATCTCTCGTAGCTTGATAAGAGTTCTCTCTCTTTTGAGAGATCGCTTCTCT 237
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 atcatctactgtga 472
      || || ||| ||| |||
Db 238 CTCTCTCACTGAGA 251
      || || ||| ||| |||

RESULT 8
US-08-819-288-1
; Sequence 1, Application US/08819288
; Patent No. 5955652
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
;

```

```

; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,288
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Beardell, Lori Y.
; REGISTRATION NUMBER: 34,293
; REFERENCE/DOCKET NUMBER: UPN-2949
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-819-288-1

Query Match 7.6%; Score 39.6; DB 2; Length 6172;
Best Local Similarity 56.0%; Pred. No. 0.017;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps

QY 339 tcccttttcgtatttcatggtttcgttgagctctttccccgcgttatgatttctct 398
Db ||||||| ||||| ||||||||||||| |||||||||
Dy 250 TCATTATCTCTCTTTCGATGAAGACTGAGCTTTCTCTCTTTCTCTCTTTCTCTCT 309
QY 399 ttgttgtctcttctgtctaaccctctcgatatatgccctctcaggtttcatctga 458
Db ||||||| ||||| ||||||||||||| |||||||||
Dy 310 CTCTATCTCTATCTCTCGTAGCTTGATAAGAGTTCTCTCTTTTGAAGATCCGTTTCTCT 369
QY 459 atcatctactgtga 472
Db || ||||| ||
Dy 370 CTCTCCTCACTGAGA 383

RESULT 9
US-09-400-348-1
; Sequence 1, Application US/09400348
; Patent No. 6355778
; GENERAL INFORMATION:
; APPLICANT: Ecker, Joseph
; APPLICANT: Alonso, Jose
; TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
; TITLE OF INVENTION: AND PATHOGENS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 6355778ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/400,348
; FILING DATE:
; CLASSIFICATION:

```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/819,288
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Beardell, Lori Y.
;; REGISTRATION NUMBER: 34,293
;; REFERENCE/DOCKET NUMBER: UPN-2949
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-568-3100
;; TELEFAX: 215-568-3439
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6172 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-09-400-348-1

Query Match
Best Local Similarity 7.6%; Score 39.6; DB 4; Length 6172;
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 339 tcccttttctgtgatttcattgcttctgagctcttcttcccccgtttatgatttctct 398
Db 250 TTTTATCTCTCTCTTCGATGGAACGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309

QY 399 tttgtctctctgtgtaaacctctctgtatataatgacctctcaggtttcatttctga 458
Db 310 CTCATCTCTCTCTCTGAGTCTGATAAGAGTTCTCTCTCTCTCTCTCTCTCTCTCT 369

QY 459 atcatctactgtga 472
Db 370 CTCCTCACTGAGA 383

RESULT 10
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; EARLIER FILING DATE: 1998-01-14
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-01-27
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc feature
; LOCATION: (1)-(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match
Best Local Similarity 7.5%; Score 38.8; DB 4; Length 289;
Matches 24; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

QY 207 cagcccaaacactccccaaagctcagccctcttgccttgaacaatcaagccgctcc 266
Db 256 CVGVCYAYAYGAYGYTYYAYCYGAYCYGAYCYGAYCYGAYCYGAYCYGAYCYGAY 197

QY 267 tggatgaactgctccgcccctctgtctcttggcagggttggagccactgtgacctgagc 326
Db 196 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 137

QY 327 gacttctatggctcccttttctgtgatttctcattgcttcttcccccgtttt 386
Db 136 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 77

QY 387 atgattttctcttttctctct-ctctgtctaaacctctctgtatataatgacctctcag 445
Db 76 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 17

QY 446 gtttcatttc 455
Db 16 AYTYYGYTY 7

Query Match
Best Local Similarity 9.6%; Score 38.8; DB 4; Length 289;
Matches 24; Conservative 107; Mismatches 118; Indels 1; Gaps 1;

QY 207 cagcccaaacactccccaaagctcagccctcttgccttgaacaatcaagccgctcc 266
Db 256 CVGVCYAYAYGAYGYTYYAYCYGAYCYGAYCYGAYCYGAYCYGAYCYGAYCYGAY 197

QY 267 tggatgaactgctccgcccctctgtctcttggcagggttggagccactgtgacctgagc 326
Db 196 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 137

QY 327 gacttctatggctcccttttctgtgatttctcattgcttcttcccccgtttt 386
Db 136 NYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 77
```



```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(176373)
; OTHER INFORMATION: n - A,T,C or G
US-09-128-155-17

Query Match          7.4%; Score 38.2; DB 3; Length 176373;
Best Local Similarity 50.3%; Pred. No. 0.22;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 329 cttctatggtccctttctgtgatttcatggtttctgagctcttttcccccgctttat 388
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144525 cttctctctctctctcttttctctctctctctctctctctctctctctctct 144584

QY 389 gattttctctctctctctctctctctctctctctctctctctctctctctctct 448
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144585 ttctttctctctctctctctctctctctctctctctctctctctctctctct 144644

QY 449 tcattctgaatactactgtgaaactattcccatgtttgcccagaagcccccgtgtct 508
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144645 ctctctctctctctctctctctctctctctctctctctctctctctctctct 144704

QY 509 tctctct 515
      |||||
Db 144705 tctcttt 144711
```

```
RESULT 15
US-08-929-329-1/c
; Sequence 1, Application US/08929329
; Patent No. 6120770
; GENERAL INFORMATION:
; APPLICANT: Adams, John H
; APPLICANT: Dalton, John P
; APPLICANT: Kappe, Stefan
; TITLE OF INVENTION: Plasmodium Proteins Useful for Preparing
; TITLE OF INVENTION: Vaccine Compositions
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Barnes & Thornburg
; STREET: 11 S Meridian
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,329
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Breen, John P
; REGISTRATION NUMBER: 38,833
; REFERENCE/DOCKET NUMBER: 835910-28685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 231-7745
; TELEFAX: (317) 231-7433
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium yoelii
```

US-08-929-329-1

Query Match 7.1%; Score 36.6; DB 3; Length 5433;  
Best Local Similarity 47.2%; Pred. No. 0.13;  
Matches 111; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 279 tcgcctctctctctctctgttgagccactgtggccctgagcgcacttctatggc 338  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3730 TTCCCTCTTCTGCTTTTTCAGCTTCCTTCAATTGCTTTTCCTTCTTCTGCTTCTTGC 3671

QY 339 tccctttctctgtgatttctcatggtttctgagctcttttcccccgctttatgatctctct 398  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3670 AGCTTCAATTATTTTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3611

QY 399 ttttctctctctctctgtaaacctctctctctctctctctctctctctctctctctctct 458  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3610 TTCTTCTGCTTTTTCAGCTTCAATTGCTTTTCCTTCTCTCTCTCTCTCTCTCTCTCT 3551

QY 459 atcatctactgtgaactattccccattgtttgccagaagcccccgtgtctctctct 513  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 3550 TTCAATTGCTTTTTCCT 3496

Search completed: June 8, 2002, 09:35:02  
Job time: 5850 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:40:18 ; Search time 369.19 Seconds  
(without alignments)  
2408.954 Million cell updates/sec

Title: US-09-503-596-8  
Perfect score: 518  
Sequence: 1 gaattccagcaggaatcagg.....ccctgggtcttctcttaga 518

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	518	100.0	518	17 AAT08982	Adipose specific e
2	518	100.0	518	21 AAA37718	Mouse AFABP enhanc
3	508.8	98.2	512	13 AAO23906	Adipose specific e
C 4	41	7.9	13076	22 AAL06214	Human reproductive
C 5	41	7.9	13076	22 AAL07355	Human reproductive
C 6	40.6	7.8	72604	20 AAZ10752	Genomic sequence o
7	39.6	7.6	6042	17 AAT09018	Arabidopsis thalia
8	39.6	7.6	6172	19 AAV57454	Arabidopsis ethyle
9	39.6	7.6	6172	22 AAD03789	Arabidopsis thalia

C 10	39.4	7.6	4590	22	AAH24065	Yeast AOB9604-asso
11	39.4	7.6	15484	22	AAK70808	Human immune/haema
12	39.2	7.6	2010	10	AAK91379	Intron 1 from huma
C 13	39	7.5	80331	22	AAC89559	Human histone deac
C 14	38.6	7.5	2118	22	AAD16391	Mouse interferon i
C 15	38.4	7.4	857	23	AAS92392	DNA encoding novel
C 16	38.2	7.4	372	20	AAZ27713	Human DNA marker C
17	38.2	7.4	5059	20	AAH84332	Stealth virus nucl
18	38.2	7.4	10620	20	AAK02996	Human IL-lra BAC c
19	38.2	7.4	14690	20	AAK22303	Human IL-lra BAC c
20	37.6	7.3	1020	22	AAF63569	Murine phosphatase
C 21	37.6	7.3	1520	22	AAI90669	Human polynucleoti
C 22	37.4	7.2	2347	23	AAS75464	DNA encoding novel
C 23	37.4	7.2	2482	23	AAS90736	DNA encoding novel
C 24	37.4	7.2	6644	22	AAF31479	ZALPHA associated
C 25	37.4	7.2	9539	22	AAK45346	Chemically pretrea
C 26	37	7.1	1466	24	AAK63068	Cell death protect
27	36.8	7.1	887	22	AAI94064	Human neuroblastom
C 28	36.8	7.1	2315	21	AAF18129	Lung cancer associ
C 29	36.8	7.1	2337	22	AAH81742	Human differentia
C 30	36.8	7.1	44453	20	AAK23519	Human kidney amino
31	36.8	7.1	49999	20	AAZ23891	Murine LOBO genomi
C 32	36.8	7.1	49999	20	AAZ23896	Murine LOBO homolo
C 33	36.6	7.1	567	22	AAH33915	Human colon cancer
C 34	36.6	7.1	5433	21	AAK99265	Plasmodium yoelii
C 35	36.4	7.0	139	24	AAK63092	Cell death protect
36	36.4	7.0	1001	21	AAH51571	Human 6PGD related
C 37	36.4	7.0	1239	23	AAK65580	DNA encoding novel
C 38	36.4	7.0	1239	23	AAK66695	DNA encoding novel
C 39	36.4	7.0	1239	23	AAK71706	DNA encoding novel
C 40	36.4	7.0	1717	23	AAK75833	DNA encoding novel
C 41	36.4	7.0	2043	23	AAK75793	DNA encoding novel
C 42	36.4	7.0	2389	18	AAK96470	MSRV-1 clone JLBcl
C 43	36.4	7.0	2389	19	AAV43161	Multiple sclerosis
C 44	36.4	7.0	5306	24	ABL32510	Human immune syste
45	36.4	7.0	34488	22	AAF97854	Human neuroblastom

## ALIGNMENTS

RESULT 1	
AAT08982	
ID AAT08982 standard; DNA; 518 BP.	
XX AAT08982;	
AC AAT08982;	
DT 11-JUL-1996 (first entry)	
DE Adipose specific enhancer element.	
XX	
DE Adipose; enhancer element; murine; adipocyte P2; gene; ap2;	
KW lipolytic; lipogenic; protein; adipose tissue; recombinant;	
KW metabolism; transgenic animal; fat tissue; homeostasis; disease;	
KW obesity; study; ds.	
XX	
OS Mus musculus.	
XX	
PN US5476926-A.	
XX	
PD 19-DEC-1995.	
XX	
PF 28-SEP-1990; 90US-0589971.	
XX	
PR 07-OCT-1991; 91US-0771022.	
PR 28-SEP-1990; 90US-0589971.	
XX	
PA (GRAV/) GRAVES R.	
PA (ROSS/) ROSS S.	
PA (SPIE/) SPIEGELMAN B M.	
XX	
PI Graves R, Ross S, Spiegelman BM;	

DE WPI; 1996-049020/05.  
XX New murine adipose-specific enhancer element - used for fat-specific  
PT expression of a heterologous gene for production of a recombinant  
PT protein  
XX  
PS Claim 2; Columns 35-36; 46pp; English.  
XX  
XX The present sequence, an adipose specific enhancer element (ASEE),  
CC is derived from the 5'-flanking region of the murine adipocyte P2  
CC (ap2) gene. A DNA sequence comprising the ASEE linked through a  
CC functional promoter to a heterologous gene, which encodes a  
CC recombinant protein with a lipolytic or lipogenic effect on  
CC adipose tissue, is capable of specifically directing the in vivo  
CC expression of the protein in adipose tissue. The DNA can be used  
CC to exert an effect on adipose tissue metabolism, and in the  
CC prodn. of transgenic animals which exhibit altered fat tissue  
CC metabolism. Such animals can be leaner or more obese, and can be  
CC used in the study of adipose homeostasis and disease states  
CC associated with adipose tissue, e.g. obesity.  
XX  
XX Sequence 518 BP; 113 A; 146 C; 108 G; 151 T; 0 other;  
SQ

Query Match 100.0%; Score 518; DB 17; Length 518;  
Best Local Similarity 100.0%; Pred. No. 3.9e-144; Indels 0; Gaps 0;  
Matches 518; Conservative 0; Mismatches 0;

QY 1 gaattccagcagaatcaggtagctggaatgcacagagccatcgattcttgcaag 60  
DB 1 gaattccagcagaatcaggtagctggaatgcacagagccatcgattcttgcaag 60  
QY 61 ccatgcgacaaagcgagaatgcacatttcacccagagagagggattggtcagcag 120  
DB 61 ccatgcgacaaagcgagaatgcacatttcacccagagagagggattggtcagcag 120  
QY 121 aagtcaccaccagagacaaatggattccacagatgctgacattgcttctactgg 180  
DB 121 aagtcaccaccagagacaaatggattccacagatgctgacattgcttctactgg 180  
QY 181 atcagagttcactagtggaagtgcacagcccaaacactccccaaagctcagcccttc 240  
DB 181 atcagagttcactagtggaagtgcacagcccaaacactccccaaagctcagcccttc 240  
QY 241 ttgctctgataacaaatcagcgcctgagcagactcttatggtccctttctgtgatttcag 360  
DB 241 ttgctctgataacaaatcagcgcctgagcagactcttatggtccctttctgtgatttcag 360  
QY 361 gttcttgagctcttttcccccgctttatgattttctcttttctctctctctgctctctct 420  
DB 361 gttcttgagctcttttcccccgctttatgattttctcttttctctctctctgctctctct 420  
QY 421 ctcctctgataacaaatcagcgcctgagcagactcttatggtccctttctgtgatttcag 480  
DB 421 ctcctctgataacaaatcagcgcctgagcagactcttatggtccctttctgtgatttcag 480  
QY 481 catgtgttgcagaagccctctggttctctctctctctctctctctctctctctctctctct 518  
DB 481 catgtgttgcagaagccctctggttctctctctctctctctctctctctctctctctctct 518

RESULT 2  
AAA37718  
ID AAA37718 standard; cDNA; 518 BP.  
XX  
AC AAA37718;  
XX  
DT 22-NOV-2000 (first entry)  
XX

DE Mouse AFABP enhancer sequence.  
XX  
KW AFABP; adipocyte fatty acid binding protein; atherosclerotic lesion;  
KW lesion formation inhibition; macrophage; adipocyte; atherosclerosis;  
KW serum cholesterol; therapy; mouse; enhancer; ss.  
XX  
OS Mus sp.  
XX  
PN WO200047734-A1.  
XX  
PD 17-AUG-2000.  
XX  
XX 11-FEB-2000; 2000WO-US03560.  
XX  
XX 12-FEB-1999; 99US-0119880.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
XX Lee M, Perrella MA, Hotamisligil GS;  
PI WPI; 2000-506094/45.  
XX  
XX Reducing expression of adipocyte fatty acid binding protein through  
PT administration of a compound is used to inhibit formation of an  
PT atherosclerotic lesion -  
XX  
PS Disclosure; Page 18; 43pp; English.  
XX  
XX This sequence is the mouse AFABP (adipocyte fatty acid binding  
CC protein) enhancer sequence. The invention relates to a method  
CC for inhibiting formation of an atherosclerotic lesion comprising  
CC administering to a mammal a compound that reduces expression of adipocyte  
CC fatty acid binding protein (AFABP). The method is used to inhibit  
CC formation of atherosclerotic lesions. The method is used to identify  
CC compounds which can be used to inhibit formation of atherosclerotic  
CC lesions through inhibition of AFABP binding to an intracellular ligand  
CC in a macrophage or adipocyte, inhibition of development of an  
CC atherosclerotic lesion, inhibition of a macrophage differentiating into  
CC a foam cell or inhibition of AFABP expression in a cell. AFABP activity  
CC may be inhibited to treat atherosclerosis or to treat individuals at risk  
CC of developing atherosclerosis. Inhibiting AFABP expression or activity  
CC reduces the development of atherosclerotic lesions despite a high level  
CC of serum cholesterol.  
XX  
SQ Sequence 518 BP; 113 A; 146 C; 108 G; 151 T; 0 other;

Query Match 100.0%; Score 518; DB 21; Length 518;  
Best Local Similarity 100.0%; Pred. No. 3.9e-144; Indels 0; Gaps 0;  
Matches 518; Conservative 0; Mismatches 0;

QY 1 gaattccagcagaatcaggtagctggaatgcacagagccatcgattcttgcaag 60  
DB 1 gaattccagcagaatcaggtagctggaatgcacagagccatcgattcttgcaag 60  
QY 61 ccatgcgacaaagcgagaatgcacatttcacccagagagagggattggtcagcag 120  
DB 61 ccatgcgacaaagcgagaatgcacatttcacccagagagagggattggtcagcag 120  
QY 121 aagtcaccaccagagacaaatggattccacagatgctgacattgcttctactgg 180  
DB 121 aagtcaccaccagagacaaatggattccacagatgctgacattgcttctactgg 180  
QY 181 atcagagttcactagtggaagtgcacagcccaaacactccccaaagctcagcccttc 240  
DB 181 atcagagttcactagtggaagtgcacagcccaaacactccccaaagctcagcccttc 240  
QY 241 ttgctctgataacaaatcagcgcctgagcagactcttatggtccctttctgttcttgga 300  
DB 241 ttgctctgataacaaatcagcgcctgagcagactcttatggtccctttctgttcttgga 300  
QY 301 ggggttgagccactgtggtcagcagactcttatggtccctttctgttcttgga 360  
DB 301 ggggttgagccactgtggtcagcagactcttatggtccctttctgttcttgga 360

Db 301 ggggtggagccacactgtggcctgagcgaactctctatgtgctccctctttctgtgattttcatg 360  
QY 361 gtttctgagctcttttccccccgcttttatgattttctctttttgtctctctcttctgtctaaac 420  
Db 361 gtttctgagctcttttccccccgcttttatgattttctctttttgtctctctcttctgtctaaac 420  
QY 421 ctctctgtatatatgcccctctcaagtttcattttctgtatcttctgaatcactactgtgaactattcc 480  
Db 421 ctctctgtatatatgcccctctcaagtttcattttctgtatcttctgaatcactactgtgaactattcc 480  
QY 481 cattgtttccagaagccccctgggtttcttccctctaga 518  
Db 481 cattgtttccagaagccccctgggtttcttccctctaga 518  
RESULT 3  
AAQ23906  
ID AAQ23906 standard; DNA; 512 BP.  
XX  
XX  
NC AAQ23906;  
23-OCT-1992 (first entry)  
DE Adipose specific enhancer ap2.  
KW SV40; small t antigen; lipolytic; liogenic; ss.  
XX  
XX Mus musculus.  
XX  
XX WO9206104-A.  
XX  
XX PD 16-APR-1992.  
XX  
XX PF 25-SEP-1991; 91WO-US06989.  
XX  
XX PR 28-SEP-1990; 90US-0589971.  
XX  
XX PA (DANC ) DANA-FARBER CANCER.  
XX  
XX PA (UNII ) UNIV OF ILLINOIS COLLEGE.  
XX  
XX PI Graves R, Ross S, Spiegelman BM;  
XX  
XX WPI; 1992-150807/18.  
XX  
XX PT Adipocyte-specific enhancers and promoters - for prodn. of  
XX PT transgenic animals exhibiting altered fat metabolism  
XX  
XX PS Disclosure; Page 77; 103pp; English.  
XX  
XX The sequence given encodes an adipose-specific enhancer, ap2, which,  
XX within the scope of the invention, can be operatively linked to a DNA  
XX coding sequence for a recombinant protein having either a lipolytic or  
XX lipogenic effect on adipose tissue. This transgene was transfected  
XX into mice. The DNA coding sequence used was a 280 bp fragment of the  
XX SV40 small t antigen splice/polyadenylation sequences.  
XX  
XX Sequence 512 BP; 110 A; 146 C; 107 G; 149 T; 0 other;  
Query Match 98.2%; Score 508.8; DB 13; Length 512;  
Best Local Similarity 99.6%; Pred. No. 2.1e-141;  
Matches 510; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 6 ccagcagggaatcaggttagctggagaatcgacagagccatgcgattcttggcgaagccatg 65  
Db 1 ccagcagggaatcaggttagctggagaatcgacagagccatgcgattcttggcgaagccatg 60  
QY 66 cgacaaaggcagaagaatgcacatttcacccagagagagaggattgatgcagcgggaagtc 125  
Db 61 cgacaaaggcagaagaatgcacatttcacccagagagagaggattgatgcagcgggaagtc 120  
QY 126 accaccacagagacaataggagttccagatgcctacatttgccttcttactggatcag 185

Db 121 accaccacagagacaataggagttccagatgcctgacatttgccttcttactggatcag 180  
QY 186 agttcactagtgaaagtgcacagcccaaacactcccccaagctcagcccttctctgccc 245  
Db 181 agttcactagtgaaagtgcacagcccaaacactcccccaagctcagcccttctctgccc 240  
QY 246 ttgtaacaatacgaagcgcctcctggatgaactgctcgcgcctctgtctcttttggcagggtt 305  
Db 241 ttgtaacaatacgaagcgcctcctggatgaactgctcgcgcctctgtctcttttggcagggtt 300  
QY 306 ggagcccaactgtgctgagcgaacttctatggtcccttttctgtgattttcatggtttc 365  
Db 301 ggagcccaactgtgctgagcgaacttctatggtcccttttctgtgattttcatggtttc 360  
QY 366 tgagctcttttcccccgctttatgattttctcttctgtctctctctgtgtaaacctcct 425  
Db 361 tgagctcttttcccccgctttatgattttctcttctgtctctctctgtgtaaacctcct 420  
QY 426 tcgtatatatgcccctctcaggtttcattttctgaatcactactgaactattccattg 485  
Db 421 tcgtatatatgcccctctcaggtttcattttctgaatcactactgaactattccattg 480  
QY 486 ttgcccagaagccccctgggtttcttccctctag 517  
Db 481 ttgcccagaagccccctgggtttcttccctctag 512  
RESULT 4  
AAL06214/C  
ID AAL06214 standard; DNA; 13076 BP.  
XX  
XX AC AAL06214;  
XX  
XX DT 21-NOV-2001 (first entry)  
XX  
XX DE Human reproductive system related antigen DNA SEQ ID NO: 8902.  
XX  
XX KW Human; reproductive system related antigen; reproductive system disorder;  
XX KW cancer; gene therapy; ds.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200155320-A2.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US01339.  
XX  
XX PR 31-JAN-2000; 2000US-0179065.  
XX PR 04-FEB-2000; 2000US-0180628.  
XX PR 24-FEB-2000; 2000US-0184664.  
XX PR 02-MAR-2000; 2000US-0186350.  
XX PR 16-MAR-2000; 2000US-0189874.  
XX PR 17-MAR-2000; 2000US-0190076.  
XX PR 18-APR-2000; 2000US-0198123.  
XX PR 19-MAY-2000; 2000US-0205515.  
XX PR 07-JUN-2000; 2000US-0209467.  
XX PR 30-JUN-2000; 2000US-0215135.  
XX PR 07-JUL-2000; 2000US-0216647.  
XX PR 11-JUL-2000; 2000US-0217487.  
XX PR 11-JUL-2000; 2000US-0217496.  
XX PR 14-JUL-2000; 2000US-0218290.  
XX PR 26-JUL-2000; 2000US-0220963.  
XX PR 26-JUL-2000; 2000US-0220964.  
XX PR 14-AUG-2000; 2000US-0224518.  
XX PR 14-AUG-2000; 2000US-0224519.  
XX PR 14-AUG-2000; 2000US-0225213.  
XX PR 14-AUG-2000; 2000US-0225214.  
XX PR 14-AUG-2000; 2000US-0225266.  
XX PR 14-AUG-2000; 2000US-0225267.  
XX PR 14-AUG-2000; 2000US-0225268.





Db	7975	ATTCTCTTCGTTCCCGTGTTCCTTCCTTTTCCTTC	7935
RESULT	5		
AAAL07355/c			
ID	AAAL07355	standard; DNA; 13076 BP.	
XX	XX		
AC	AAAL07355;		
XX	XX		
DT	21-NOV-2001	(first entry)	
XX	XX		
DE	Human reproductive system related antigen DNA SEQ ID NO: 10043.		
XX	XX		
KW	Human; reproductive system related antigen; reproductive system disorder;		
KW	cancer; gene therapy; ds.		
OS	Homo sapiens.		
XX	XX		
PN	WO200155320-A2.		
XX	XX		
	02-AUG-2001.		
	17-JAN-2001;	2001WO-US01339.	
XX	XX		
PR	31-JAN-2000;	2000US-0179065.	08-SEP-2000; 2000US-0232080.
PR	04-FEB-2000;	2000US-0180628.	08-SEP-2000; 2000US-0232081.
PR	24-FEB-2000;	2000US-0184664.	12-SEP-2000; 2000US-0231968.
PR	02-MAR-2000;	2000US-0186350.	PR 14-SEP-2000; 2000US-0232397.
PR	16-MAR-2000;	2000US-0189874.	PR 14-SEP-2000; 2000US-0232398.
PR	17-MAR-2000;	2000US-0190076.	PR 14-SEP-2000; 2000US-0232399.
PR	18-APR-2000;	2000US-0198123.	PR 14-SEP-2000; 2000US-0232400.
PR	19-MAY-2000;	2000US-0205515.	PR 14-SEP-2000; 2000US-0232401.
PR	07-JUN-2000;	2000US-0209467.	PR 14-SEP-2000; 2000US-0233063.
PR	28-JUN-2000;	2000US-0214886.	PR 14-SEP-2000; 2000US-0233064.
PR	30-JUN-2000;	2000US-0215135.	PR 14-SEP-2000; 2000US-0233065.
PR	07-JUL-2000;	2000US-0216647.	PR 21-SEP-2000; 2000US-0234223.
PR	11-JUL-2000;	2000US-0216880.	PR 21-SEP-2000; 2000US-0234224.
PR	11-JUL-2000;	2000US-0217487.	PR 25-SEP-2000; 2000US-0234274.
PR	11-JUL-2000;	2000US-0217496.	PR 25-SEP-2000; 2000US-0234997.
PR	14-JUL-2000;	2000US-0218290.	PR 25-SEP-2000; 2000US-0234998.
PR	26-JUL-2000;	2000US-0220963.	PR 26-SEP-2000; 2000US-0235484.
PR	26-JUL-2000;	2000US-0220964.	PR 27-SEP-2000; 2000US-0235834.
PR	14-AUG-2000;	2000US-0224518.	PR 27-SEP-2000; 2000US-0235836.
PR	14-AUG-2000;	2000US-0224519.	PR 29-SEP-2000; 2000US-0236327.
PR	14-AUG-2000;	2000US-0225211.	PR 29-SEP-2000; 2000US-0236367.
PR	14-AUG-2000;	2000US-0225214.	PR 29-SEP-2000; 2000US-0236368.
PR	14-AUG-2000;	2000US-0225266.	PR 29-SEP-2000; 2000US-0236369.
PR	14-AUG-2000;	2000US-0225267.	PR 02-OCT-2000; 2000US-0236370.
PR	14-AUG-2000;	2000US-0225268.	PR 02-OCT-2000; 2000US-0236802.
PR	14-AUG-2000;	2000US-0225270.	PR 02-OCT-2000; 2000US-0237037.
PR	14-AUG-2000;	2000US-0225447.	PR 02-OCT-2000; 2000US-0237038.
PR	14-AUG-2000;	2000US-0225757.	PR 02-OCT-2000; 2000US-0237039.
PR	14-AUG-2000;	2000US-0225758.	PR 02-OCT-2000; 2000US-0237040.
PR	14-AUG-2000;	2000US-0225759.	PR 13-OCT-2000; 2000US-0239335.
PR	18-AUG-2000;	2000US-0226279.	PR 13-OCT-2000; 2000US-0239337.
PR	22-AUG-2000;	2000US-0226681.	PR 20-OCT-2000; 2000US-0240960.
PR	22-AUG-2000;	2000US-0226686.	PR 20-OCT-2000; 2000US-0241221.
PR	22-AUG-2000;	2000US-0227182.	PR 20-OCT-2000; 2000US-0241785.
PR	23-AUG-2000;	2000US-0227009.	PR 20-OCT-2000; 2000US-0241786.
PR	30-AUG-2000;	2000US-0228924.	PR 20-OCT-2000; 2000US-0241808.
PR	01-SEP-2000;	2000US-0229287.	PR 20-OCT-2000; 2000US-0241809.
PR	01-SEP-2000;	2000US-0229343.	PR 01-NOV-2000; 2000US-0241826.
PR	01-SEP-2000;	2000US-0229344.	PR 01-NOV-2000; 2000US-0241826.
PR	01-SEP-2000;	2000US-0229345.	PR 08-NOV-2000; 2000US-0246474.
PR	05-SEP-2000;	2000US-0229509.	PR 08-NOV-2000; 2000US-0246475.
PR	05-SEP-2000;	2000US-0229513.	PR 08-NOV-2000; 2000US-0246476.
PR	06-SEP-2000;	2000US-0230437.	PR 08-NOV-2000; 2000US-0246477.
PR	06-SEP-2000;	2000US-0230438.	PR 08-NOV-2000; 2000US-0246478.
PR	08-SEP-2000;	2000US-0231242.	PR 08-NOV-2000; 2000US-0246523.
PR	08-SEP-2000;	2000US-0231243.	PR 08-NOV-2000; 2000US-0246524.
PR	08-SEP-2000;	2000US-0231244.	PR 08-NOV-2000; 2000US-0246525.
PR	08-SEP-2000;	2000US-0231244.	PR 08-NOV-2000; 2000US-0246526.
PR	08-SEP-2000;	2000US-0231413.	PR 08-NOV-2000; 2000US-0246527.
PR	08-SEP-2000;	2000US-0231414.	PR 08-NOV-2000; 2000US-0246528.
PR	08-SEP-2000;	2000US-0231414.	PR 08-NOV-2000; 2000US-0246532.
PR	08-SEP-2000;	2000US-0249208.	PR 08-NOV-2000; 2000US-0246609.
PR	08-SEP-2000;	2000US-0249209.	PR 08-NOV-2000; 2000US-0246610.
PR	08-SEP-2000;	2000US-0249210.	PR 08-NOV-2000; 2000US-0246611.
PR	08-SEP-2000;	2000US-0249211.	PR 08-NOV-2000; 2000US-0246613.
PR	08-SEP-2000;	2000US-0249212.	PR 17-NOV-2000; 2000US-0249207.
PR	08-SEP-2000;	2000US-0249213.	PR 17-NOV-2000; 2000US-0249208.
PR	08-SEP-2000;	2000US-0249214.	PR 17-NOV-2000; 2000US-0249209.
PR	08-SEP-2000;	2000US-0249215.	PR 17-NOV-2000; 2000US-0249210.
PR	08-SEP-2000;	2000US-0249216.	PR 17-NOV-2000; 2000US-0249211.
PR	08-SEP-2000;	2000US-0249217.	PR 17-NOV-2000; 2000US-0249212.
PR	08-SEP-2000;	2000US-0249218.	PR 17-NOV-2000; 2000US-0249213.
PR	08-SEP-2000;	2000US-0249244.	PR 17-NOV-2000; 2000US-0249214.
PR	08-SEP-2000;	2000US-0249245.	PR 17-NOV-2000; 2000US-0249215.
PR	08-SEP-2000;	2000US-0249264.	PR 17-NOV-2000; 2000US-0249216.
PR	08-SEP-2000;	2000US-0249265.	PR 17-NOV-2000; 2000US-0249217.
PR	08-SEP-2000;	2000US-0249297.	PR 17-NOV-2000; 2000US-0249218.
PR	08-SEP-2000;	2000US-0249299.	PR 17-NOV-2000; 2000US-0249219.









PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235835.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483426/52.  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and  
XX metastasis -  
XX Disclosure; SEQ ID NO 25620; 3071pp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
XX amino acid sequences given in AAM91921. (I) have cytostatic  
CC









**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 08:43:34 ; Search time 2258.96 Seconds  
(without alignments)  
4798.644 Million cell updates/sec

Title: us-09-503-596-8  
Perfect score: 518  
Sequence: 1 gaattccacaggaatcagg.....ccctgtttcttctcttaga 518

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pi.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	518	100.0	518	6	I16725	I16725 Sequence 1
2	518	100.0	518	10	MUSAB2A	M84651 Mus musculu
3	82.6	15.9	165948	9	AC018616	AC018616 Homo sapi
4	82.6	15.9	174404	2	AC013542	AC013542 Homo sapi
5	74.2	14.3	7218	6	I66494	I66494 Sequence 14
6	71.6	13.8	507	11	G32180	G32180 STS 8.3640F
7	48.6	9.4	190491	2	AC091260	AC091260 Mus muscu
8	47.4	9.2	164520	2	AC020738	AC020738 Homo sapi
9	47	9.1	173381	2	AC107311	AC107311 Homo sapi
10	46.2	8.9	174742	9	AC068102	AC068102 Homo sapi
11	46.2	8.9	182150	9	HSG278N14	AL109654 Human DNA
12	45.8	8.8	56509	2	AC099974	AC099974 Mus muscu
13	45.8	8.8	24963	2	AL663072	AL663072 Mus muscu
14	45.4	8.7	121788	2	AC097569	AC097569 Rattus no
15	45.2	8.7	184056	2	AC024551	AC024551 Homo sapi
16	44.8	8.6	208613	2	AC093464	AC093464 Mus muscu
17	44.6	8.6	188892	2	AC091327	AC091327 Mus muscu
18	44.4	8.6	219424	2	AC110011	AC110011 Homo sapi
19	44.4	8.6	309401	2	AC016584	AC016584 Homo sapi
20	44.2	8.5	86514	9	AL136117	AL136117 Human DNA
21	44.2	8.5	86701	9	AL353654	AL353654 Human DNA
22	44	8.5	144141	2	AC098932	AC098932 Homo sapi
23	43.8	8.5	194968	2	AC083757	AC083757 Homo sapi
24	43.8	8.5	246262	2	AL669825	AL669825 Mus muscu
25	43.4	8.4	165058	2	AC105361	AC105361 Homo sapi
26	43.2	8.3	178100	9	AC009080	AC009080 Homo sapi
27	43.2	8.3	184529	9	AC009156	AC009156 Homo sapi
28	43.2	8.3	192060	2	AC106728	AC106728 Homo sapi
29	43.2	8.3	192247	2	AC013497	AC013497 Homo sapi
30	43.2	8.3	203156	2	AC026303	AC026303 Homo sapi
31	43	8.3	209842	2	AC074335	AC074335 Mus muscu
32	42.8	8.3	114956	9	AC002072	AC002072 Human PAC
33	42.8	8.3	132327	9	AL158209	AL158209 Human DNA
34	42.8	8.3	173978	2	AC020649	AC020649 Homo sapi
35	42.8	8.3	178220	2	AC027482	AC027482 Homo sapi
36	42.8	8.3	192724	2	AC073825	AC073825 Mus muscu
37	42.8	8.3	200000	2	AC005769	AC005769 Homo sapi
38	42.8	8.3	209062	2	AC005284	AC005284 Homo sapi
39	42.6	8.2	41437	9	AC006134	AC006134 Homo sapi
40	42.6	8.2	87421	2	AC098622	AC098622 Rattus no
41	42.6	8.2	127615	9	AL356461	AL356461 Human DNA
42	42.6	8.2	134539	9	HS24M15	Z94055 Human DNA s
43	42.6	8.2	146451	2	AC019309	AC019309 Homo sapi
44	42.6	8.2	169513	9	AP002777	AP002777 Homo sapi
45	42.6	8.2	170758	9	AC004965	AC004965 Homo sapi

ALIGNMENTS

RESULT	1	I16725	Sequence 1 from patent US 5476926.	518 bp	DNA	linear	PAT 03-APR-1996
LOCUS	I16725	I16725	I16725				
DEFINITION	I16725	I16725	I16725				
ACCESSION	I16725	I16725	I16725				
VERSION	I16725.1	GI:1251633					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 518)						
AUTHORS	Spiegelman, B.M., Graves, R. and Ross, S.						
TITLE	Adipocyte-specific DNA sequences and use thereof in the production of transgenic animals exhibiting altered fat tissue metabolism						
JOURNAL	Patent: US 5476926-A 1 19-DEC-1995;						
FEATURES	Location/Qualifiers						
source	1.518						
BASE COUNT	113 a	146 c	108 g	151 t			
ORIGIN							
Query Match		100.0%	Score 518;	DB 6;	Length 518;		

Best Local Similarity 100.0%; Pred. No. 7.8e-140;  
Matches 518; Conservative 0; Mismatches 0;

Best local similarity 100.0%, Freq. NO: 7.0e-140,  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	gaattccagcaggaatcaagtagctagctgaggagaatcgacagagccatcgcatctcttggcaag	60
Db	1	GAATTCAGCAGGAATCAGTAGCTGAGAAATCGACAGAGCCATCGGATTCCTGGCAAG	60
Qy	61	ccatgcgcaaaaggcagaaatgcacatttcaccacagagagaaggattgatgtcagcagag	120
Db	61	CCATGCGCAAAAGGCAGAAATGCACATTTCACCACAGAGAAAGGATTCATGTCACACAG	120
Qy	121	aagtcaaccacccagagagcaaatggagtttcccagatgcctgcagatttgctctcttactgg	180
Db	121	AAGTCACACCCACAGAGACAAATGGAGTTCCCAAGTGCCTGCATTTGCCCTCTTACTGG	180
Qy	181	atcagagttcactagtggaagtgcacagcccaaacactcccccaagctcagccctcc	240
Db	181	ATCAGAGTTCACTAGTGGAGAGTGTACAGGCCAAACACTCCCCCAAGCTCAGCCCTCC	240
Qy	241	ttgccttgtaacaatcaagccgctcctggatgaactgctcgcctctctctcttcttggca	300
Db	241	TTGCCTTGTAAACAATCAAGCCGCTCGTGGATGAACACTGCCCGCCCTCTGTCTCTTGGCA	300
Qy	301	gggttgagccactgtggcctgagcgactctatagctccctcttctctgtatttcctg	360
Db	301	GGGTTGAGGCCCACTGTGGCCTGAGGCGACTTCATGGCTCCCTTTTCTGTGATTTTCATG	360
Qy	361	gtttctgagctcttttcccccggttttatgatattctctcttttgcctctctcttgcataac	420
Db	361	GTTTCTGAGCTCTTTTCCCCCGGTTATGATTTTCTCTTTTGTCTCTCTCTTGCTAAAC	420
Qy	421	ctcctcgatatatgccctctcaggtttcattcttgaatcactactactgaactattcc	480
Db	421	CTCCTTCGTATATATGCCCCCTCAGGTTTCATTTCTGAATCACTACTGTGAACATATCC	480
Qy	481	cattgtttgcagaagcccctggttctctctctctaga	518
Db	481	CATTGTTTGCAGAAGCCCTCGTCTTCTCCTCTAGA	518

RESULT	2	
MUSAP2A		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		

FEATURES	source	Location/Qualifiers
		1..518
		/organism="Mus musculus"
		/strain="Swiss"
		/db_xref="taxon:10090"
gene		1..518
		/gene="ap2"
enhancer		1..518
		/gene="ap2"
BASE COUNT	113 a 146 c 108 g 151 t	
ORIGIN		

```
Query Match      100.0%; Score 518; DB 10; Length 518;
Best Local Similarity 100.0%; Pred. No. 7.8e-140;
Matches 518; Conservative 0; Mismatches 0; Indels 0
```

Qy	1	gaattccagcagaataatcagtagtctgtagaataatccacagagccatcgattcttggcaag	60
Db	1	GAATTCAGCAGGAATCAGTAGTCTGGAGAAATGCACAGAGCCATCGATTCTTTGGCAAG	60
Qy	61	ccatgcgacaaaagcagaataatcacatttcaccacagagagaaggtatgatgtccagcag	120
Db	61	CCATGCGACAAAGCGCAAAATGCACATTTCACCCAGAGAGAGGGATTGATGTCACAGG	120
Qy	121	aagtcaaccaccagagagcaaatggagtttccagatgctctgacatttgcccttcttactgg	180
Db	121	AAGTCAACACCCAGAGAGCAAAATGGAGTTCCCAAGATGCCTGACATTTCGCTTCTTACTGG	180
Qy	181	atcacagattcactagtggaagtgtcacagcccaaacctcccccaagctcagcccttcc	240
Db	181	ATCAGAGTTCACTAGTGGAAAGTGTACAGGCCCAAAACACTCCCCCAAGCTCAGCCCTTC	240
Qy	241	ttgccttgtaacaatacaagccgctcctggatgaactgctcgccctctgtctctttggcca	300
Db	241	TTGCCTTGTAAACAATCAAGCCGCTCTGGAGTGAACCTGCTCGGCCCTCTGTCTCTTTGGCA	300
Qy	301	ggattgagcccactgtggcctgaagcaactctatagctccctcttctgtgatttcattg	360
Db	301	GGGTTGAGGCCACTGTTGGCCTGAGGCACTCTATGGTCCCTTTTCTGTGATTTTCATG	360
Qy	361	gtttctgagctctttccccccgctttatgatattctctttttgtctctctctctgctaaac	420
Db	361	GTTTCTCAGCTCTTTTCCCCCGCTTTATGATTTTCTCTTTTGTCTCTCTCTGCTGAAC	420
Qy	421	ctccttcgatatatgccctctcaggtttcattctcgaatcatctactactgaactattcc	480
Db	421	CTCCTTCGATATATGCCCCCTCAGGGTTTCATTTCGAATCATCTACTGTGAACTATTC	480
Qy	481	cattgtttgcagaagcccctggttctctctctaga	518
Db	481	CATTGTTTGCAGAAGCCCCTGTTCTCTCTCTCTAGA	518

### RESULT 3

AC018616	165948 bp	DNA	linear	PRI 30-MAY-2001
AC018616	Human sapiens chromosome 8, clone RP11-15714, complete sequence.			
AC018616				
AC018616.5	GI:12429085			
HTG.				
human.				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
1 (bases 1 to 165948)				
Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
Homo sapiens chromosome 8, clone RP11-15714				
Unpublished				
2 (bases 1 to 165948)				
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,				
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,				
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G., Castle,A.,				
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,				
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Feneater,J.,				
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,				
Gradya,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,				
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,				
Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K.,				
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,				
Meldrum,T., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,				
O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N.,				
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,				
Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,				
Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A.,				
Vassiliou,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.				
and Zody,M.				
Direct Submission				
Submitted (14-DEC-1999)				
Whitehead Institute/MIT Center for Genome				
TITLE				
JOURNAL				

TITLE  
JOURNAL

Submitted (14-DEC-1999) Whitehead Institute/MIT Center for Genome  
Direct Submission  
and Zou/M.





[illegible]

[illegible]









Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,  
Severy, P., Sougnès, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
Direct Submission  
Submitted (29-MAY-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

**REFERENCE**  
4 (bases 1 to 174742)

**AUTHORS**  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, D., Boukhgalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heard, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
Direct Submission  
Submitted (10-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

**COMMENT**  
On Dec 10, 2001 this sequence version replaced gi:14211833.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L7452

Center clone name: 333\_M\_8

-----

# FEATURES

## source

Location/Qualifiers

1..174742

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-333M8"

/clone\_lib="RPC1-11 Human Male BAC"

complement(892..1004)

/rpt\_family="LIMC5"

complement(1375..1949)

/rpt\_family="L2"

complement(2211..2524)

/rpt\_family="AluSp"

complement(2784..3385)

/rpt\_family="LIP8"

complement(4551..4679)

/rpt\_family="FLAM\_A"

complement(4743..4827)

/rpt\_family="L2"

complement(5125..5185)

/rpt\_family="MIR3"

complement(5238..5702)

/rpt\_family="L2"

5867..5891

repeat\_region  
complement(5892..5936)  
/rpt\_family="LIMB4"

repeat\_region  
complement(6391..6538)  
/rpt\_family="MLT1H1"

repeat\_region  
complement(6633..6749)  
/rpt\_family="MLT1J"

repeat\_region  
complement(6879..7162)  
/rpt\_family="L2"

repeat\_region  
complement(7163..7252)  
/rpt\_family="LIP8A16"

repeat\_region  
complement(7253..7751)  
/rpt\_family="L2"

repeat\_region  
8414..8884  
/rpt\_family="LIM4"

repeat\_region  
9604..9641  
/rpt\_family="AT-rich"

repeat\_region  
complement(9701..9951)  
/rpt\_family="LIME4A"

repeat\_region  
complement(10715..11084)  
/rpt\_family="LIMA9"

repeat\_region  
11085..12580  
/rpt\_family="LIMAS4"

repeat\_region  
12581..12608  
/rpt\_family="(TG)n"

repeat\_region  
12609..12654  
/rpt\_family="LIMAS4"

repeat\_region  
12655..12957  
/rpt\_family="AluX"

repeat\_region  
12958..13013  
/rpt\_family="LIMAS4"

repeat\_region  
complement(14726..14770)  
/rpt\_family="L2"

repeat\_region  
complement(15971..16132)  
/rpt\_family="MIR"

repeat\_region  
complement(16224..16379)  
/rpt\_family="LIME4A"

repeat\_region  
16517..16594  
/rpt\_family="MADE1"

repeat\_region  
16662..16735  
/rpt\_family="(TA)n"

repeat\_region  
16744..16773  
/rpt\_family="(CA)n"

repeat\_region  
complement(16851..18314)  
/rpt\_family="LIMC4a"

repeat\_region  
complement(18387..18913)  
/rpt\_family="LIMC4a"

repeat\_region  
19375..19435  
/rpt\_family="AT-rich"

repeat\_region  
19559..19600  
/rpt\_family="(TTA)n"

repeat\_region  
complement(19856..20415)  
/rpt\_family="LIMC5"

repeat\_region  
complement(20770..21207)  
/rpt\_family="HAL1b"

repeat\_region  
21232..21506  
/rpt\_family="AluX"

repeat\_region  
21964..22012  
/rpt\_family="CT-rich"

repeat\_region  
22149..22185  
/rpt\_family="T-rich"

repeat\_region  
22190..22215  
/rpt\_family="(TTCA)n"

repeat\_region  
22735..22766  
/rpt\_family="(T)n"

repeat\_region  
complement(23045..23147)  
/rpt\_family="LIMC4a"

repeat\_region  
23148..23624  
/rpt\_family="MLT1D"

repeat\_region  
complement(23625..24440)  
/rpt\_family="LIMC4a"

repeat\_region  
complement(24441..24733)  
/rpt\_family="AluSg"

```
Query Match          8.9%; Score 46.2; DB 9; Length 174742;
Best Local Similarity 61.0%; Pred. No. 0.022;
Matches 75; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 343 tttctgtgatttcaggtttctgagctcttttcccccgcgttatgattttctctttt 402
      ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 45238 TTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT 45297
      TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT

QY 403 gtctctctctgtaaacctctctgttatatgcctctcaggtttcattcttctgaatca 462
      ||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 45298 TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 45357
      TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT

QY 463 tct 465
      ||
Db 45358 CCT 45360

RESULT 11
LOCUS HSG278N14
DEFINITION Human DNA sequence from clone GSI-278N14 on chromosome xq27.3-28.
          Contains STSS and GSSs, complete sequence.
ACCESSION AL109654
VERSION AL109654.22 GI:6599103
KEYWORDS HTG
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 182150)
Direct Submission
Submitted (11-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 17, 1999 this sequence version replaced gi:6562113.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
GSI-278N14 is
from the library Genome_Systems_Release1 VECTOR: pBelOBAC11
This sequence is the entire insert of clone GSI-278N14. This
sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX.
Location/Qualifiers
1..182150
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q27.3-28"
/clone="GSI-278N14"
/clone.lib="Genome_Systems_Release1"
185..224
/note="8 copies 5 mer aaaa 85% conserved"
repeat_region
1233..1332
/note="MIR repeat: matches 136. .247 of consensus"
repeat_region
1336..1589
/note="RHEIC repeat: matches 1. .371 of consensus"
```

```
repeat_region
1605..1644
/note="L1M1 repeat: matches -1386. -.1120 of consensus"
repeat_region
2016..9603
/note="MIR repeat: matches 108. .148 of consensus"
repeat_region
9710..9833
/note="L1PA14 repeat: matches -1421. .6149 of consensus"
repeat_region
10304..10520
/note="MIR repeat: matches 59. .183 of consensus"
misc_feature
10598..10813
/note="L2 repeat: matches 2498. .2710 of consensus"
repeat_region
10779..10822
/note="match: GSS: Em:AQ065649"
misc_feature
11126..11251
/note="22 copies 2 mer ac 79% conserved"
misc_feature
11128..11637
/note="match: STS: Em:L24630"
repeat_region
11916..12093
/note="match: GSS: Em:AQ440005"
repeat_region
12217..13077
/note="MIR repeat: matches 81. .262 of consensus"
repeat_region
13078..13440
/note="L1MB5 repeat: matches 5335. .6166 of consensus"
repeat_region
13441..13494
/note="THE1B repeat: matches 1. .364 of consensus"
misc_feature
13954..14493
/note="L1MB5 repeat: matches 5281. .5335 of consensus"
repeat_region
15311..15479
/note="match: GSS: Em:AQ116486"
repeat_region
15481..15561
/note="MER34 repeat: matches 6. .175 of consensus"
repeat_region
16475..16918
/note="MER20 repeat: matches 140. .218 of consensus"
repeat_region
16938..17039
/note="L1MA1 repeat: matches 5757. .6198 of consensus"
repeat_region
17050..17899
/note="L1PA9 repeat: matches 6062. .6163 of consensus"
repeat_region
17903..21450
/note="L1PA10 repeat: matches 5273. .6158 of consensus"
repeat_region
21526..21878
/note="L1P3 repeat: matches 391. .3762 of consensus"
repeat_region
21888..23168
/note="L1P3 repeat: matches 10. .374 of consensus"
repeat_region
23141..25156
/note="L1M2 repeat: matches 4477. .5766 of consensus"
repeat_region
25123..27030
/note="L1M1 repeat: matches 916. .2926 of consensus"
misc_feature
complement(27218..27658)
/note="L1M1 repeat: matches -1179. .740 of consensus"
repeat_region
27304..27364
/note="match: GSS: Em:AQ177664"
repeat_region
27473..27581
/note="MER82 repeat: matches 45. .105 of consensus"
repeat_region
27659..27867
/note="MIR repeat: matches 31. .138 of consensus"
repeat_region
28139..28628
/note="MER82 repeat: matches 444. .652 of consensus"
repeat_region
30972..31194
/note="MLT1D repeat: matches 1. .505 of consensus"
repeat_region
31636..31742
/note="L1TR33 repeat: matches 14. .286 of consensus"
repeat_region
31743..31791
/note="L1PA2 repeat: matches 6040. .6146 of consensus"
repeat_region
32343..32384
/note="L1PA3 repeat: matches 5969. .6017 of consensus"
misc_feature
33730..34237
/note="21 copies 2 mer aa 81% conserved"
repeat_region
33853..34068
/note="match: GSS: Em:AQ773243"
repeat_region
34829..34858
/note="MER3 repeat: matches 1. .209 of consensus"
repeat_region
34861..35229
/note="15 copies 2 mer ca 93% conserved"
repeat_region
/note="RHEIC repeat: matches 1. .371 of consensus"
```

repeat\_region 35230..36830 /note="THE1C-internal repeat: matches 1..1580 of consensus"  
repeat\_region 36831..37194 /note="THE1C repeat: matches 1..371 of consensus"  
repeat\_region 37609..37913 /note="THE1B repeat: matches 1..335 of consensus"  
repeat\_region 37931..38024 /note="MIR repeat: matches 167..262 of consensus"  
repeat\_region 38730..38840 /note="MER5A repeat: matches 43..159 of consensus"  
repeat\_region 39030..39383 /note="THE1B repeat: matches 1..364 of consensus"  
repeat\_region 39384..40959 /note="THE1B-INTERNAL repeat: matches 1..1580 of consensus"  
repeat\_region 40960..41000 /note="THE1B repeat: matches 314..354 of consensus"  
repeat\_region 41011..41078 /note="LIP1A7 repeat: matches 6058..6125 of consensus"  
repeat\_region 41109..41432 /note="THE1B repeat: matches 1..335 of consensus"  
repeat\_region 41585..44290 /note="HERVL repeat: matches 2050..4761 of consensus"  
repeat\_region 44285..44951 /note="HERVL repeat: matches 1247..1914 of consensus"  
repeat\_region 44954..45257 /note="LIP1A3 repeat: matches 5861..6148 of consensus"  
repeat\_region 45281..45675 /note="LIP1A repeat: matches 3624..4024 of consensus"  
misc\_feature complement(45585..46171) /note="match: GSS: Em:AQ382752"  
misc\_feature complement(45678..46175) /note="match: GSS: Em:AQ382427"  
repeat\_region 46063..46282 /note="LIP1A10 repeat: matches 5936..6153 of consensus"  
misc\_feature 46183..46654 /note="match: GSS: Em:AQ607587"  
misc\_feature 46188..46452 /note="match: STS: Em:L42721"  
repeat\_region 46421..48233 /note="LIP3 repeat: matches 392..1986 of consensus"  
repeat\_region 48227..50485 /note="LIP1A8 repeat: matches 3880..6159 of consensus"  
repeat\_region 50504..50723 /note="LIME1 repeat: matches 3331..3552 of consensus"  
misc\_feature 50712..51416 /note="match: GSS: Em:AQ308841"  
repeat\_region 50724..51095 /note="MLT1D repeat: matches 113..505 of consensus"  
repeat\_region 51144..53304 /note="LIME1 repeat: matches 3557..5758 of consensus"  
repeat\_region 53308..53392 /note="MLT1A1 repeat: matches 2..86 of consensus"  
repeat\_region 53393..53701 /note="AluY repeat: matches 1..309 of consensus"  
repeat\_region 53730..54011 /note="MLT1A1 repeat: matches 73..354 of consensus"  
repeat\_region 54015..54445 /note="LIME1 repeat: matches 5754..6172 of consensus"  
repeat\_region 54757..55479 /note="LIP1A4 repeat: matches 5393..6142 of consensus"  
repeat\_region 55769..55921 /note="MER5A repeat: matches 7..183 of consensus"  
repeat\_region 56385..57268 /note="LIME3 repeat: matches 5244..6162 of consensus"  
repeat\_region 57351..57561 /note="AluSg repeat: matches 1..209 of consensus"  
repeat\_region 57844..57877 /note="LTR9 repeat: matches 593..625 of consensus"  
repeat\_region 57878..58043 /note="MER61E repeat: matches 346..502 of consensus"  
repeat\_region 58044..58060

/note="LTR9 repeat: matches 404..593 of consensus"  
Query Match 8.9%; Score 46.2; DB 9; Length 182150;  
Best Local Similarity 61.0%; Pred. No. 0.022;  
Matches 75; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 344 ttctgtgatttcacgttcgtgagctctttcccccgcgttatgatcttctctttt 403  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 122262 TTCTCTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122321  
QY 404 tctctctctgtctaaacctctctcgatatatgcctctctcaggttctcttgaatcat 463  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 122322 TTCTCTTCT 122381  
QY 464 cta 466  
||  
Db 122382 GTA 122384  
AC099974 56509 bp DNA linear HTG 22-NOV-2001  
Mus musculus clone RP23-22J3, LOW-PASS SEQUENCE SAMPLING.  
AC099974  
AC099974.1 GI:17047340  
HTG: HTGS\_PHASE0.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP23-22J3  
Unpublished  
2 (bases 1 to 56509)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,N., Matthews,C.,  
McLean,C., Macdonald,P., Major,J., Marquis,R., McPheeters,R., Meldrim,J.,  
McCarthy,M., McEwan,P., McKernan,K., Murphy,T., Naylor,J., Nguyen,C.,  
Meneus,L., Mihova,T., Mienga,V., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Norbu,C., Norman,C.H., Phunkhang,P., Pierre,R., Rise,C., Rogov,P.,  
Oliver,J., Peterson,K., Rieback,M., Riley,R., Santos,R., Schauback,R.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,N.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: LI3695  
Center clone name: 22\_J3  
-----  
\* NOTE: This record contains 70 individual

\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 711: contig of 711 bp in length  
\* 712 811: gap of 100 bp  
\* 812 1512: contig of 701 bp in length  
\* 1513 1612: gap of 100 bp  
\* 1613 2331: contig of 719 bp in length  
\* 2332 2431: gap of 100 bp  
\* 2432 3145: contig of 714 bp in length  
\* 3146 3245: gap of 100 bp  
\* 3246 3958: contig of 713 bp in length  
\* 3959 4058: gap of 100 bp  
\* 4059 4781: contig of 723 bp in length  
\* 4782 4881: gap of 100 bp  
\* 4882 5584: contig of 703 bp in length  
\* 5585 5684: gap of 100 bp  
\* 5685 6396: contig of 712 bp in length  
\* 6397 6496: gap of 100 bp  
\* 6497 7281: contig of 785 bp in length  
\* 7282 7381: gap of 100 bp  
\* 7382 8058: contig of 677 bp in length  
\* 8059 8158: gap of 100 bp  
\* 8159 8868: contig of 710 bp in length  
\* 8869 8968: gap of 100 bp  
\* 8969 9682: contig of 714 bp in length  
\* 9683 9782: gap of 100 bp  
\* 9783 10496: contig of 714 bp in length  
\* 10497 10596: gap of 100 bp  
\* 10597 11305: contig of 709 bp in length  
\* 11306 11405: gap of 100 bp  
\* 11406 12093: contig of 688 bp in length  
\* 12094 12193: gap of 100 bp  
\* 12194 12887: contig of 694 bp in length  
\* 12888 12987: gap of 100 bp  
\* 12988 13703: contig of 716 bp in length  
\* 13704 13803: gap of 100 bp  
\* 13804 14517: contig of 714 bp in length  
\* 14518 14617: gap of 100 bp  
\* 14618 15337: contig of 720 bp in length  
\* 15338 15437: gap of 100 bp  
\* 15438 16140: contig of 703 bp in length  
\* 16141 16240: gap of 100 bp  
\* 16241 16950: contig of 710 bp in length  
\* 16951 17050: gap of 100 bp  
\* 17051 17768: contig of 718 bp in length  
\* 17769 17868: gap of 100 bp  
\* 17869 18584: contig of 716 bp in length  
\* 18585 18684: gap of 100 bp  
\* 18685 19395: contig of 711 bp in length  
\* 19396 19495: gap of 100 bp  
\* 19496 20179: contig of 684 bp in length  
\* 20180 20279: gap of 100 bp  
\* 20280 20965: contig of 686 bp in length  
\* 20966 21065: gap of 100 bp  
\* 21066 21774: contig of 709 bp in length  
\* 21775 21874: gap of 100 bp  
\* 21875 22572: contig of 698 bp in length  
\* 22573 22672: gap of 100 bp  
\* 22673 23396: contig of 724 bp in length  
\* 23397 23496: gap of 100 bp  
\* 23497 24326: contig of 830 bp in length  
\* 24327 24426: gap of 100 bp  
\* 24427 25144: contig of 718 bp in length  
\* 25145 25244: gap of 100 bp  
\* 25245 25945: contig of 701 bp in length

25946 26045: gap of 100 bp  
26046 26750: contig of 705 bp in length  
26751 26850: gap of 100 bp  
26851 27562: contig of 712 bp in length  
27563 27662: gap of 100 bp  
27663 28307: contig of 645 bp in length  
28308 28407: gap of 100 bp  
28408 29118: contig of 711 bp in length  
29119 29218: gap of 100 bp  
29219 29929: contig of 711 bp in length  
29930 30029: gap of 100 bp  
30030 30754: contig of 725 bp in length  
30755 30854: gap of 100 bp  
30855 31577: contig of 723 bp in length  
31578 31677: gap of 100 bp  
31678 32390: contig of 713 bp in length  
32391 32490: gap of 100 bp  
32491 33199: contig of 709 bp in length  
33200 33299: gap of 100 bp  
33300 34011: contig of 712 bp in length  
34012 34111: gap of 100 bp  
34112 34816: contig of 705 bp in length  
34817 34916: gap of 100 bp  
34917 35608: contig of 692 bp in length  
35609 35708: gap of 100 bp  
35709 36375: contig of 667 bp in length  
36376 36475: gap of 100 bp  
36476 37172: contig of 697 bp in length  
37173 37272: gap of 100 bp  
37273 37985: contig of 713 bp in length  
37986 38085: gap of 100 bp  
38086 38769: contig of 684 bp in length  
38770 38869: gap of 100 bp  
38870 39570: contig of 701 bp in length  
39571 39670: gap of 100 bp  
39671 40376: contig of 706 bp in length  
40377 40476: gap of 100 bp  
40477 41186: contig of 710 bp in length  
41187 41286: gap of 100 bp  
41287 42000: contig of 714 bp in length  
42001 42100: gap of 100 bp  
42101 42799: contig of 699 bp in length  
42800 42899: gap of 100 bp  
42900 43620: contig of 721 bp in length  
43621 43720: gap of 100 bp  
43721 44444: contig of 724 bp in length  
44445 44544: gap of 100 bp  
44545 45257: contig of 713 bp in length  
45258 45357: gap of 100 bp  
45358 46065: contig of 708 bp in length  
46066 46165: gap of 100 bp  
46166 46863: contig of 698 bp in length  
46864 46963: gap of 100 bp  
46964 47676: contig of 713 bp in length  
47677 47776: gap of 100 bp  
47777 48474: contig of 698 bp in length  
48475 48574: gap of 100 bp  
48575 49272: contig of 698 bp in length  
49273 49372: gap of 100 bp  
49373 50085: contig of 713 bp in length  
50086 50185: gap of 100 bp  
50186 50897: contig of 712 bp in length  
50898 50997: gap of 100 bp  
50998 51682: contig of 685 bp in length  
51683 51782: gap of 100 bp  
51783 52537: contig of 755 bp in length  
52538 52637: gap of 100 bp  
52638 53339: contig of 702 bp in length  
53340 53439: gap of 100 bp  
53440 54135: contig of 696 bp in length  
54136 54235: gap of 100 bp  
54236 54913: contig of 678 bp in length  
54914 55013: gap of 100 bp

```

* 55014 55702: contig of 689 bp in length
* 55703 55802: gap of 100 bp

Query Match      8.8%; Score 45.8; DB 2; Length 56509;
Best Local Similarity 53.7%; Pred. No. 0.026;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 342 cttttctgtgatttcattggttctctgagctcttttccccgcgtttatgatttctctttt 401
Db 6090 cttttctgttctttcttttcttttcttttcttttcttttcttttcttttcttttctttt 6031

QY 402 tgtctctctcttctaaacctctctgtatatatgcctctcagggttcattctgaatc 461
Db 6030 tcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttctttt 5971

QY 462 atctactgtgaactattccattgtttgcccagaagccccctggttcttctctctaga 518
Db 5970 tcttttctgagatgggtttctctgtgtgagccctggcgatcctagaaactactctctaga 5914

RESULT 13
AL663072      248963 bp      DNA      linear      HTG 01-FEB-2002
LOCUS      Mus musculus chromosome X clone RP23-22J3, *** SEQUENCING IN
DEFINITION      PROGRESS ***, in unordered pieces.
ACCESSION      AL663072
VERSION      AL663072.8 GI:18491423
KEYWORDS      HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1 (sites)
JOURNAL      Howden, P.
COMMENT      Direct Submission
Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:18476984.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bm22J3
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 248349 bases at least Q40
Consensus quality: 248604 bases at least Q30
Consensus quality: 248705 bases at least Q20
Insert size: 248763; sum-of-contigs
Insert size: 248868; 2.0% error; agarose-fp
Quality coverage: 9.38x in Q20 bases; sum-of-contigs Quality
coverage: 9.46x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
source      Location/Qualifiers
1. 248963
   /organism="Mus musculus"
   /db_xref="taxon:10090"
   /chromosome="X"
   /clone="RP23-22J3"
   /clone_lib="RPCI-23"
   1. 129645
   /note="assembly fragment:04132
   fragment_chain:1"
   129746..125400
   /note="assembly_fragment:05296

```

```

fragment_chain:1"
225501..248963
/note="assembly fragment:01912
fragment_chain:1"
BASE COUNT 76690 a 51750 c 50560 g 69759 t 204 others
ORIGIN

```

```

Query Match      8.8%; Score 45.8; DB 2; Length 248963;
Best Local Similarity 53.7%; Pred. No. 0.029;
Matches 95; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 342 cttttctgtgatttcattggttctctgagctcttttccccgcgtttatgatttctctttt 401
Db 175161 ctttttcttttcttttcttttcttttcttttcttttcttttcttttcttttctttt 175220

QY 402 tgtctctctcttctaaacctctctgtatatatgcctctcagggttcattctgaatc 461
Db 175221 tcttttcttttcttttcttttcttttcttttcttttcttttcttttcttttctttt 175280

QY 462 atctactgtgaactattccattgtttgcccagaagccccctggttcttctctctaga 518
Db 175281 tcttttctgagatgggtttctctgtgtgagccctggcgatcctagaaactactctctaga 175337

```

```

RESULT 14
AC097569      121788 bp      DNA      linear      HTG 20-DEC-2001
LOCUS      Rattus norvegicus clone CH230-67E3, *** SEQUENCING IN PROGRESS ***,
DEFINITION      65 unordered pieces.
ACCESSION      AC097569
VERSION      AC097569.3 GI:17974367
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 121788)

```

```

REFERENCE
AUTHORS      Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
      Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbara, J.,
      Benton, J.P., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J.,
      Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
      Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
      Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
      Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
      Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
      Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
      Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
      Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
      Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
      Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
      Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
      Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
      Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
      Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
      Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
      Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
      Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
      Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
      Loulsegod, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
      Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
      Martinez, E., Massey, E., Mahoney, E., McLeod, M.P., Meador, M.,
      Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
      Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Mohabbat, K.,
      Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
      Oghu, M., Okwuonu, G., Oragunye, N., Owiedo, R., Pace, A., Payton, B.,
      Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
      Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojokan, I., Rolfe, M.,
      Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
      Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,
      Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
      Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,

```



Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 JOURNAL

Unpublished  
 2 (bases 1 to 121788)  
 Worley, K.C.

Submitted (19-OCR-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:17062605.

COMMENT

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GITZ  
 Center clone name: CH230-67E3  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to  
 findPhrapList  
 Consensus quality: 93320 bases at least Q40  
 Consensus quality: 102603 bases at least Q30  
 Consensus quality: 109123 bases at least Q20  
 Estimated insert size: 63172; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-fp estimation  
 Quality coverage: 0.8x in Q20 bases; sum-of-contigs estimation

-----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 65 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 6540: contig of 6540 bp in length  
 6641 6640: gap of unknown length  
 6641 10202: contig of 3562 bp in length  
 10203 10302: gap of unknown length  
 10303 13913: contig of 3611 bp in length  
 13914 14013: gap of unknown length  
 14014 16759: contig of 2746 bp in length  
 16760 16859: gap of unknown length  
 16860 20400: contig of 3541 bp in length  
 20401 20500: gap of unknown length  
 20501 22351: contig of 1851 bp in length  
 22352 22451: gap of unknown length  
 22452 25002: contig of 2551 bp in length  
 25003 25102: gap of unknown length  
 25103 27203: contig of 2101 bp in length  
 27204 27303: gap of unknown length  
 27304 30014: contig of 2711 bp in length  
 30015 32339: contig of 2225 bp in length  
 32340 32439: gap of unknown length  
 32440 34411: contig of 1972 bp in length  
 34412 34511: gap of unknown length  
 34512 37116: contig of 2605 bp in length  
 37117 37216: gap of unknown length  
 37217 39412: contig of 2196 bp in length  
 39413 39512: gap of unknown length  
 39513 41595: contig of 2083 bp in length  
 41596 44635: gap of unknown length  
 44636 44735: gap of unknown length  
 44736 47198: contig of 2463 bp in length  
 44736

47199 47298: gap of unknown length  
 47299 48779: contig of 1361 bp in length  
 48680 48779: gap of unknown length  
 48780 50621: contig of 1842 bp in length  
 50622 52803: contig of 2082 bp in length  
 52804 52903: gap of unknown length  
 52904 54588: contig of 1685 bp in length  
 54589 54688: gap of unknown length  
 54689 55941: contig of 1253 bp in length  
 55942 56042: gap of unknown length  
 56043 57498: contig of 1457 bp in length  
 57499 57598: gap of unknown length  
 57599 59373: contig of 1775 bp in length  
 59374 59474: gap of unknown length  
 59475 60760: contig of 1287 bp in length  
 60761 60860: gap of unknown length  
 60861 61944: contig of 1084 bp in length  
 61945 62044: gap of unknown length  
 62045 63497: contig of 1452 bp in length  
 63498 63597: gap of unknown length  
 63598 65187: gap of unknown length  
 65188 66267: contig of 1340 bp in length  
 66268 66726: gap of unknown length  
 66727 68371: contig of 1645 bp in length  
 68372 68471: gap of unknown length  
 68472 69962: contig of 1491 bp in length  
 69963 70063: gap of unknown length  
 70064 71864: contig of 1802 bp in length  
 71865 71965: gap of unknown length  
 71966 73645: contig of 1681 bp in length  
 73646 73746: gap of unknown length  
 73747 74900: contig of 1055 bp in length  
 74901 76755: gap of unknown length  
 76756 76854: gap of unknown length  
 76855 77870: contig of 1016 bp in length  
 77871 77970: gap of unknown length  
 77971 79214: contig of 1144 bp in length  
 79215 80266: contig of 1052 bp in length  
 80267 81995: contig of 1629 bp in length  
 81996 82095: gap of unknown length  
 82096 83503: contig of 1408 bp in length  
 83504 84837: gap of unknown length  
 84838 84937: gap of unknown length  
 84938 86327: contig of 1390 bp in length  
 86328 86427: gap of unknown length  
 86428 88067: contig of 1640 bp in length  
 88068 88167: gap of unknown length  
 88168 89675: gap of unknown length  
 89676 90812: contig of 1037 bp in length  
 90813 90912: gap of unknown length  
 90913 92400: contig of 1488 bp in length  
 92401 92501: gap of unknown length  
 92502 93625: contig of 1125 bp in length  
 93626 93725: gap of unknown length  
 93726 95247: contig of 1522 bp in length  
 95248 95348: gap of unknown length  
 95349 96716: contig of 1369 bp in length  
 96717 96817: gap of unknown length  
 96818 98331: contig of 1815 bp in length  
 98332 98731: gap of unknown length  
 98732 100983: contig of 2252 bp in length  
 100984 101083: gap of unknown length  
 101084 102243: contig of 1160 bp in length  
 102244 103948: gap of unknown length  
 103949 104048: contig of 1605 bp in length  
 104049 104048: gap of unknown length



```
* 89119 89218: gap of 100 bp
* 89219 100341: contig of 11123 bp in length
* 100342 100441: gap of 100 bp
* 100442 109823: contig of 9382 bp in length
* 109824 109923: gap of 100 bp
* 109924 118272: contig of 8349 bp in length
* 118273 118372: gap of 100 bp
* 118373 128451: contig of 10079 bp in length
* 128452 128552: gap of 100 bp
* 128552 138311: contig of 9760 bp in length
* 138312 138411: gap of 100 bp
* 138412 151032: contig of 12621 bp in length
* 151033 151132: gap of 100 bp
* 151133 168498: contig of 17366 bp in length
* 168499 168598: gap of 100 bp
* 168599 184056: contig of 15458 bp in length.
FEATURES
    source
        1. 184056
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="4"
            /map="4"
            /clone="RP11-585K20"
            /clone_lib="RPC1-11 Human Male BAC"
            1. 1057
                /note="assembly_fragment"
            1158. 2655
                /note="assembly_fragment"
            2756. 4142
                /note="assembly_fragment"
            4243. 5444
                /note="assembly_fragment"
            5545. 6973
                /note="assembly_fragment"
            7074. 8743
                /note="assembly_fragment"
            8844. 11607
                /note="assembly_fragment"
            11708. 14116
                /note="assembly_fragment"
            clone_end:r7
            vector_side:right
            14217. 15094
                /note="assembly_fragment"
            clone_end:SP6
            vector_side:right
            15195. 18623
                /note="assembly_fragment"
            18724. 21108
                /note="assembly_fragment"
            21209. 24350
                /note="assembly_fragment"
            24451. 29418
                /note="assembly_fragment"
            29519. 33760
                /note="assembly_fragment"
            33861. 39779
                /note="assembly_fragment"
            39880. 45932
                /note="assembly_fragment"
            46033. 52475
                /note="assembly_fragment"
            52576. 57348
                /note="assembly_fragment"
            57449. 63623
                /note="assembly_fragment"
            63724. 71408
                /note="assembly_fragment"
            71509. 80078
                /note="assembly_fragment"
            80179. 89118
                /note="assembly_fragment"
            89219. 100341
```

```
misc_feature
    /note="assembly_fragment"
    100442. 109823
misc_feature
    /note="assembly_fragment"
    109924. 118272
misc_feature
    /note="assembly_fragment"
    118373. 128451
misc_feature
    /note="assembly_fragment"
    128552. 138311
misc_feature
    /note="assembly_fragment"
    138412. 151032
misc_feature
    /note="assembly_fragment"
    151133. 168498
Query Match      8.7%; Score 45.2; DB 2; Length 184056;
Best Local Similarity 34.8%; Pred. NO. 0.042;
Matches 119; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
QY 115 acagagaagtcaccacccagagagcaaatggagttcccccagatgcctgacatttgccttct 174
    |||||
Db 138484 ACAGGAAGCAGAGAGCTGGGACTGGAATGGATTGATCCGAGCCCTCACTTTGGATTGCT 138425
QY 175 tactggatcagagttcactagtggaagtgtcacagcccaaacacactcccccagctcagc 234
    |||
Db 138424 TCCTCTACCACACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 138365
QY 235 ccttcctgccttgtaacaatcaagccgcctcctgagatgcctgcctcctctctctct 294
    |||
Db 138364 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 138305
QY 295 ttggcagggttgagcccaactgtgcccctgagcgacttctatggctccctttctgtgatt 354
    |||
Db 138304 TCCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 138245
QY 355 ttcatgggttctgagctcttttcccccgcctttatgatttctctctctctctctctctg 414
    |||||
Db 138244 TCTCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 138185
QY 415 ctaaacctcctcgtatatatgcctcctcaggtttcattct 456
    |||||
Db 138184 CTTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 138143
```

Search completed: June 8, 2002, 08:48:23  
Job time: 10206 sec

**THIS PAGE BLANK (USPTO)**